Profit Maximization in Bilateral Trade against a Smooth Adversary

Simone Di Gregorio
Department of Computer,
Control and Management Engineering
Sapienza University of Rome
simone.digregorio@uniroma1.it

Paul Dütting
Google Research
duetting@google.com

Federico Fusco
Department of Computer,
Control and Management Engineering
Sapienza University of Rome
federico.fusco@uniroma1.it

Chris Schwiegelshohn
Department of Computer Science
Aarhus University
cschwiegelshohn@gmail.com

Abstract

Bilateral trade models the task of intermediating between two strategic agents, a seller and a buyer, who wish to trade a good. We study this problem from the perspective of a profit-maximizing broker within an online learning framework, where the agents' valuations are generated by a smooth adversary. We devise a learning algorithm that guarantees a $\tilde{O}(\sqrt{T})$ regret bound, which is tight in the time horizon T up to poly-logarithmic factors. This matches the minimax rate for the stochastic i.i.d. case, and is also well separated from the adversarial setting, where sublinear-regret is unattainable. By extending the strong regret guarantees from the i.i.d. case to the smooth adversary, we significantly broaden the scope of settings where such a fast rate is achievable, while closing an important gap in the regret landscape of this fundamental economic problem. To overcome the challenges posed by this adversary, we leverage a continuity property of smooth instances and combine this with a hierarchical net-construction of the broker's action space, which is analyzed via algorithmic chaining. We showcase the applicability of these techniques by deriving a similarly tight $\tilde{O}(\sqrt{T})$ regret bound for a related mechanism design model: the joint ads problem.

1 Introduction

In bilateral trade, a broker intermediates between two agents—a seller and a buyer—who are interested in trading a good [Vickrey, 1961, Myerson and Satterthwaite, 1983]. Both agents are characterized by a private valuation for the item, and their goal is to maximize their own utility. The broker, on the other hand, aims at maximizing some economic metric, while ensuring that (i) agents report truthfully their valuations (*dominant-strategy incentive compatibility*) and (ii) the utility for participating in the mechanism of each agent is non-negative (*individual rationality*).

In this paper, we focus on *profit-maximization*, a canonical objective in economics which naturally arises in the applications (e.g., all e-commerce platforms that connect sellers to buyers). In particular, we investigate profit maximization in bilateral trade through the lens of online learning, with the goal of designing an algorithm that minimizes the regret with respect to the best mechanism in hindsight. For this problem, Di Gregorio et al. [2025] provide two surprising results: while the best mechanism can be learned *very* quickly in the stochastic setting where the agents' valuations are

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

drawn i.i.d. from an unknown but fixed distribution, *no algorithm* can achieve sublinear regret in the standard adversarial setting. This leaves open the natural question of whether it is possible to perform learning *beyond* the stationary i.i.d. setting. Following a consolidated approach [Rakhlin et al., 2011, Haghtalab et al., 2020, Block et al., 2022, Haghtalab et al., 2024] rooted in *beyond-worst-case-analysis* [Roughgarden, 2020], we investigate the learnability of the problem against the σ-smooth adversary. Such an adversary can choose different distributions over valuations at different time steps, but is constrained to distributions that are "not too peaked" (see Definition 2).

A characterizing feature of our learning task is the complexity of the broker's action space. A mechanism for bilateral trade is specified by two functions—an allocation and a payment rule—mapping bids to allocation and payments. Standard mechanism design arguments (see Appendix C.1) allow for a simple geometric characterization of the generic $M \in \mathcal{M}$. Its allocation rule corresponds to a subdivision of the bids space into two regions: one in which the trade happens and one where this is not the case. This partition has to be closed in the "up-left" direction, so that increasing the buyer's bid or decreasing the seller's does not decrease the likelihood of observing a trade. The payments are then given by suitable projections of the valuation pair on the boundary of the allocation region, à la Myerson [1981]. While this characterization collapses the learning problem to tracking the best allocation region, the mechanism space is still non-parametric! This separates us from the majority of the online learning literature on mechanism design, where the learning objective is intrinsically parametric, e.g., the best price(s) [Kleinberg and Leighton, 2003, Cesa-Bianchi et al., 2024c,b], the best reserve price(s) [Cesa-Bianchi et al., 2015, Roughgarden and Wang, 2019], or the best bid [Feng et al., 2018, Cesa-Bianchi et al., 2024a]. We refer to Appendix A for more related work.

1.1 Our Result

Our first contribution is the design of the learning algorithm HIER-MECH, which enjoys a regret rate of $O(1/\sigma \cdot \sqrt{T} \log T)$ for profit maximization in bilateral trade against a σ-smooth adversary (Theorem 1). This is, up to lower order terms, the same bound achievable in the i.i.d. setting, while no sublinear algorithm can exist in the (non-smooth) adversarial setting [Di Gregorio et al., 2025].

We then formally prove a reduction from the joint ads problem to profit-maximization in bilateral trade (Theorem 2). Joint ads is a mechanism design setting where two agents *collaborate* to purchase a non-excludable good, and has been studied in the online learning setting in Aggarwal et al. [2024]. The reduction implies a lower bound of $\Omega(\sqrt{T})$ on the regret for bilateral trade (Corollary 1), matching the performance of HIER-MECH. Therefore, we characterize the time-horizon dependence of the minimax regret rate against the σ-smooth adversary, up to poly-log terms. This settles an open question in Di Gregorio et al. [2025]. Finally, the reduction allows us to adapt HIER-MECH to the joint ads setting, maintaining the same (tight) regret rate. This improves on the previous regret upper bound of $O(1/\sigma \cdot T^{2/3})$, settling an open question in Aggarwal et al. [2024] (see Corollary 2).

In combination, our results thus show that, from an information-theoretic perspective, the σ-smooth adversary is "as easy as" the i.i.d. setting. Interestingly, the $\sqrt{T}$ rate algorithms for these two data-generation models are obtained with very different techniques and arguments (see Section 1.2 for more details). The strong statistical learning guarantees for the smooth adversarial setting reinforce the effectiveness of smoothness as a meaningful way of overcoming adversarial impossibility results in economic problems [Cesa-Bianchi et al., 2024c,b,a, Durvasula et al., 2023]. A remaining discrepancy is that the running time of HIER-MECH scales as $\exp(\sqrt{T})$, while the i.i.d. setting admits a poly-time algorithm. Closing this computational gap and tightening the regret dependency on the smoothness parameter σ remain the primary open questions in this space.

1.2 Technical Challenges

The main challenge of profit-maximization in bilateral trade is the non-parametric nature of the action space, and the special "projection-like" nature of the payments, which induce the profit function.

The Adversary is Non-Stationary. In the stochastic i.i.d. setting, Di Gregorio et al. [2025] address the learning problem by proposing a modified version of follow-the-leader, arguing via *probabilistic chaining* [Talagrand, 2021] that (i) a data and distribution dependent *finite* family $\mathcal{M}'$ of mechanisms has its profit estimated very quickly and (ii) all mechanisms outside $\mathcal{M}'$ can be ruled out quickly as candidate optima. Such an approach, however, is doomed to fail in any non-stationary setting, like

the one we are studying. Indeed, for any mechanism M, past profits can be completely uncorrelated with future ones, so that one can never "rule out" mechanisms which are suboptimal up to that point.

Exploiting Lipschitzness. Smoothness makes the expected-profit function Lipschitz in the choice of the mechanism (for a proper distance). Following the standard approach in the Lipschitz bandits literature [Kleinberg et al., 2019], we can then restrict our attention to a suitable metric cover and run some off-the-shelf algorithm on it. A natural candidate is the family of ε-grid mechanisms $\mathcal{M}_\varepsilon$, whose allocation region is made up of tiles of the uniform ε-grid of the unit square $[0,1]^2$ (see Figure 2). Any mechanism $M \in \mathcal{M}$ can be approximated by some $M_\varepsilon \in \mathcal{M}_\varepsilon$, so that the overall expected profit of M_ε is at least that of M, up to an additive ε/σ term (as in e.g., Lemma 5.1 of Aggarwal et al. [2024]). Unfortunately, applying directly a no-regret algorithm for adversarial prediction with experts on $\mathcal{M}_\varepsilon$ (e.g., HEDGE) yields a suboptimal rate (as $\log|\mathcal{M}_\varepsilon| \approx 1/\varepsilon$, see Lemma 1):

$$R_T \lesssim \underbrace{\varepsilon/\sigma \cdot T}_{\text{Discretization}} + \underbrace{\sqrt{T \log|\mathcal{M}_\varepsilon|}}_{\text{HEDGE guarantee in } \mathcal{M}_\varepsilon} \lesssim \frac{\varepsilon}{\sigma} T + \sqrt{\frac{T}{\varepsilon}} \approx \frac{1}{\sigma} T^{2/3}. \qquad \text{(For } \varepsilon \approx T^{-1/3}\text{)}$$

Algorithmic Chaining. To get $\sqrt{T}$, we exploit the specific structure of the mechanism space $\mathcal{M}$. In particular, we adopt *algorithmic chaining* [Cesa-Bianchi et al., 2017] to mimic *in an adversarial setting* what probabilistic chaining does in the stochastic model. At a high level, the idea is to maintain a hierarchical decomposition of the mechanism space composed of suitable *nets* at different precisions, and to construct an "hedge-like" distribution via a random walk on such structure. This allows to get a "Dudley-type" bound [e.g., Chapter 8 in Vershynin, 2026], improving on the standard $\sqrt{T \log N}$ expert bound (with N being the number of experts). While our approach is inspired by the HIERHEDGE algorithm [Appendix H of Cesa-Bianchi et al., 2017], there are several important differences and new challenges that we need to overcome. Indeed, algorithmic chaining is an algorithmic *template*, which still needs to be instantiated with the "right" choices of losses/rewards and hierarchical decomposition of the action space, and such choice is naturally problem-dependent. In particular, while in Cesa-Bianchi et al. [2017] the authors require *deterministic* L^∞ net guarantees on the hypothesis class, we know that such nets cannot exist for our problem (see Appendix B for details). We overcome this issue by providing probabilistic net guarantees, obtained via a novel L^1 bound in expectation (Lemma 2) for ε-grid mechanisms, which relies on smoothness and bilateral trade-specific features.

Joint Ads Reduction. While *some* relation between bilateral trade and joint ads is already hinted at in previous work [Di Gregorio et al., 2025], we derive a formal reduction from the latter to the former. This allows us to apply our algorithm to joint ads (implying a tight $\tilde{O}(1/\sigma \cdot \sqrt{T})$ regret rate); moreover, we extend the $\Omega(\sqrt{T})$ joint ads lower bound of Aggarwal et al. [2024] to bilateral trade.

2 Model and Preliminaries

We study the online task of learning the profit maximizing mechanism for bilateral trade, when the agents' valuations are generated by a smooth adversary. We focus on the family $\mathcal{M}$ of all the mechanisms which enforce dominant strategy incentive compatibility (DSIC) and ex-post individual rationality (IR). These two standard properties ensure that (i) the seller and the buyer declare their private valuations truthfully and (ii) they do not suffer negative utility in participating in the mechanism. We refer to Appendix C.1 for more details on the game-theoretic aspects of the problem.

The seller valuation v_s and the buyer one v_b belong to $[0,1]$, so a mechanism M is characterized by an allocation region $A \subseteq [0,1]^2$, and pricing rules $p, q : [0,1]^2 \to [0,1]$. The trade happens if and only if the agents' valuations (v_s, v_b) belong to A, and payments are made according to p and q. That is, the seller receives $p(v_s, v_b)$, while the buyer pays $q(v_s, v_b)$. Without loss of generality, we require that $p(v_s, v_b) = q(v_s, v_b) = 0$ whenever $(v_s, v_b) \notin A$, i.e., there is no trade. The profit earned by the broker running mechanism M is the difference between the payment the mechanism collects from the buyer and the payment that it makes to the seller: $\mathsf{profit}(M, (v_s, v_b)) = q(v_s, v_b) - p(v_s, v_b)$.

Characterization of DSIC and IR Mechanisms. Standard mechanism design arguments prescribe that any DSIC and IR mechanism has a distinctive structure: the payments are uniquely induced by the allocation region which, in turn, has to respect a monotonicity property.

Definition 1 (Monotone Allocation Region & Myerson Payments)**.** *An allocation region $A \subseteq [0,1]^2$ is monotone if for any $x = (x_1, x_2) \in A$ and $y = (y_1, y_2) \in [0,1]^2$ with $x_1 \geq y_1$ and $x_2 \leq y_2$, it holds that $y \in A$. A mechanism is monotone if its allocation region is monotone. Given a monotone allocation region, then the associated Myerson payments are defined as follows:*

$$p(v_s, v_b) = \mathbb{1}_{\{(v_s,v_b)\in A\}} \cdot \max\{x \in [0,1] \mid (x, v_b) \in A\} \qquad \forall\ (v_s, v_b) \in [0,1]^2$$
$$q(v_s, v_b) = \mathbb{1}_{\{(v_s,v_b)\in A\}} \cdot \min\{y \in [0,1] \mid (v_s, y) \in A\} \qquad \forall\ (v_s, v_b) \in [0,1]^2$$

In words, for an allocation region to be monotone it should be "closed" in a north-west direction. That is, if a point (v_s, v_b) is in A then any point (v'_s, v_b) with $v'_s \leq v_s$ and any point (v_s, v'_b) with $v'_b \geq v_b$ should be in A as well. The profit of $(v_s, v_b) \in A$, in turn, corresponds to the "south" projection minus the "east" one onto the allocation boundary. For the above definition to be well-posed, we require (without loss of generality) that all allocation regions are topologically closed.

Proposition 1. *A mechanism for bilateral trade is dominant-strategy incentive compatible and individually rational if and only if its allocation region is monotone and uses Myerson payments.*

Repeated Bilateral Trade. We study the following problem: At each time step t, a new seller-buyer pair arrives, with private valuations $v^t = (v_s^t, v_b^t) \in [0,1]^2$. The learner/broker proposes a mechanism $M^t \in \mathcal{M}$, and the agents bid (truthfully because of the DSIC property). The trade occurs according to M^t and the bids of the agents, which then leave forever. The broker earns $\mathsf{profit}(M^t, v^t)$, the difference between the price paid *by* the buyer and that paid *to* the seller. For simplicity, $\mathsf{profit}_t(\cdot)$ denotes the profit induced by v^t at time t. Formally, we have the following:

The Learning Protocol

for time $t = 1, 2, \ldots$ **do**
  a new pair of agents arrives with private valuations $v^t = (v_s^t, v_b^t) \in [0,1]^2$
  the learner declares a mechanism $M^t \in \mathcal{M}$
  the trade takes place according to mechanism M^t and the bids
  the learner gains $\mathsf{profit}(M^t, v^t) = \mathsf{profit}_t(M^t) \in [-1, 1]$

Smooth Distributions. We assume that the private valuations of the agents are generated by a σ-smooth adversary. More precisely, an (oblivious) adversary generates up-front a sequence $\mathcal{S}$ of independent σ-smooth random variables. A random variable is smooth if it is characterized by a density function that is not "too peaked", as formalized in the following definition.

Definition 2 (Smooth Distributions)**.** *A distribution $\mathcal{D}$ over $[0,1]^2$ is said to be σ-smooth, for some $\sigma \in (0,1]$, if the following inequality holds for any Borel set A in $[0,1]^2$: $\mathbb{P}_{v\sim\mathcal{D}}(v \in A) \leq \frac{\mathcal{L}(A)}{\sigma}$, where $\mathcal{L}$ is the Lebesgue measure. A random variable is σ-smooth if its distribution is σ-smooth.*

Regret. The regret suffered by a learning algorithm $\mathcal{A}$ onto an adversarial sequence of σ-smooth random variables $\mathcal{S}$ is defined as

$$R_T(\mathcal{A}, \mathcal{S}) = \sup_{M\in\mathcal{M}} \sum_{t=1}^{T} \mathbb{E}\left[\mathsf{profit}(M, v^t) - \mathsf{profit}(M^t, v^t)\right], \tag{1}$$

where the expectation is with respect to the random valuations $v^1, v^2, \ldots$ and possibly on the internal randomness of the algorithm. The regret of $\mathcal{A}$ is defined by its performance on the worst sequence of σ-smooth distributions: $R_T(\mathcal{A}) = \sup_{\mathcal{S}} R_T(\mathcal{A}, \mathcal{S})$. The typical goal is to achieve sublinear worst-case regret. The algorithm's average profit converges to that of the benchmark, on any input.

3 The Grid Mechanisms

A central ingredient in our algorithmic construction rests on the concept of *grid* mechanisms, i.e., mechanisms whose allocation region can be described in terms of a uniform grid on the $[0,1]^2$ square. While this family is already used in previous work [Aggarwal et al., 2024], we are the first to prove the L^1-net guarantee needed in the chaining analysis. The family of ε-grid mechanisms is parameterized by a precision ε which we assume, for convenience, to be a power of 2 ($\varepsilon = 2^{-h}$,

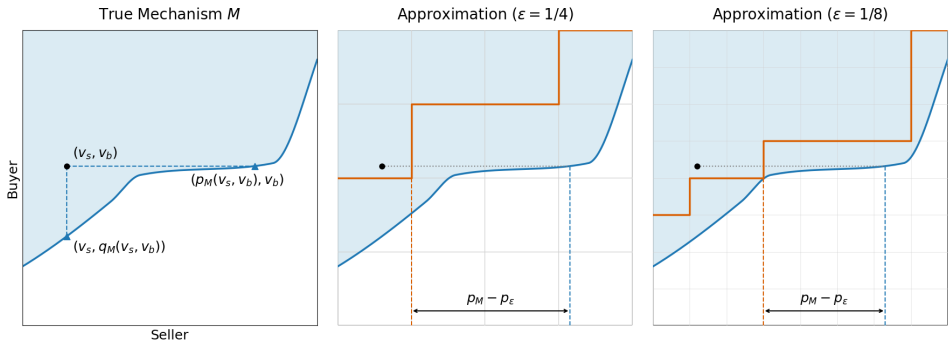

Figure 1: A monotone mechanism and two of its approximations. Notice the seller payment error.

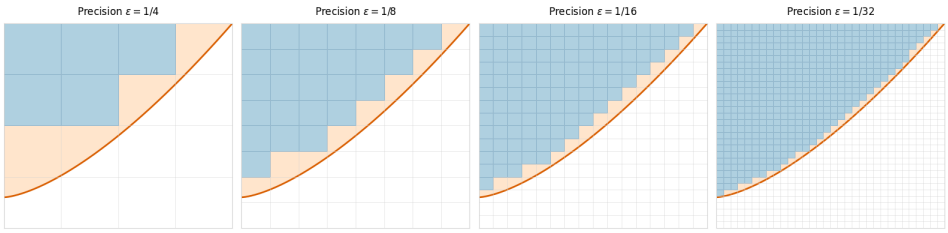

Figure 2: An illustration of the approximating map defined in Definition 4, for different values of ε. The approximation of the allocation region is in blue, the original mechanism has orange boundary.

for some $h \in \mathbb{N}$). Consider the uniform grid V_ε of step-size ε of $[0,1]^2$, i.e., the points of the form $(k\varepsilon, j\varepsilon)$ for k and j in $0, \dots, 1/\varepsilon$. The grid naturally partitions the $[0,1]^2$ square into square "tiles" of the form $[k\varepsilon, (k+1)\varepsilon] \times [j\varepsilon, (j+1)\varepsilon]$, we denote with $\mathcal{G}_\varepsilon$ the set of such tiles.

Definition 3 (Grid Mechanisms)**.** *The family of ε-grid mechanisms $\mathcal{M}_\varepsilon \subseteq \mathcal{M}$ is composed of all the mechanisms whose (monotone) allocation region is given by union of tiles in $\mathcal{G}_\varepsilon$.*

We slightly overload the definitions and require all allocation regions of mechanisms in $\mathcal{M}$ (and thus also of grid mechanisms) to contain the segments $(0,0)-(0,1)$ and $(0,1)-(1,1)$. This assumption is compatible with monotonicity and is without loss of generality (by smoothness, adding the segments does not affect expected profit). The boundary of the allocation region of a generic ε-grid mechanism is then given by the union of segments joining points in the underlying uniform grid. This allows to bound $|\mathcal{M}_\varepsilon|$ via a simple combinatorial argument, deferred to Appendix E.

Lemma 1. *The cardinality of ε-grid mechanisms respects $2^{1/\varepsilon} \leq |\mathcal{M}_\varepsilon| \leq (2e)^{1/\varepsilon}$.*

The smoothness allows to nicely approximate the profit of any mechanism via ε-grid mechanism, via some sort of ε-grid "inner" approximation: given a mechanism M, just consider the union of all the tiles in $\mathcal{G}_\varepsilon$ which are fully contained in the allocation region of M (see Figures 1 and 2).

Definition 4 (ε-approximations)**.** *For any $\varepsilon > 0$ and $M \in \mathcal{M}$ with allocation region A, we define $\phi_\varepsilon(M) \in \mathcal{M}_\varepsilon$ as the mechanism whose allocation region is the union of all $P \in \mathcal{G}_\varepsilon$ s.t. $P \subseteq A$.*

If no tile is fully contained in the allocation region of a given mechanism, then the corresponding approximating mechanism is $M_\emptyset$, which only allocates in the union of the segments $(0,0)-(0,1)$ and $(0,1)-(1,1)$ (i.e., with zero probability). Clearly, due to the smoothness assumption, such mechanism has 0 expected profit. For a generic mechanism M, the associated ε-approximated mechanism is a good approximation to the expected profit of M. We have the following "L^1-bound".

Lemma 2. *Let v be a σ-smooth random variable on $[0,1]^2$. For every $\varepsilon > 0$ we have the following inequality: $\mathbb{E}\left[|\mathsf{profit}(M, v) - \mathsf{profit}(\phi_\varepsilon(M), v)|\right] \leq 6 \cdot \varepsilon/\sigma$.*

Proof. Consider a generic mechanism M with allocation region A_M and payments p_M, q_M, and its ε-grid approximation $\phi_\varepsilon(M)$, with allocation region A_ε and payments $p_\varepsilon, q_\varepsilon$. We can decompose

the distance in profit between the two mechanisms depending on whether the realized valuation falls inside A_ε or in $A_M \setminus A_\varepsilon$. In the latter case, the profit of $\phi_\varepsilon(M)$ is zero outside its allocation region:

$$\mathbb{E}\left[|\mathsf{profit}(M,v) - \mathsf{profit}(\phi_\varepsilon(M),v)|\mathbb{1}_{\{v\in A_M\setminus A_\varepsilon\}}\right] \leq \mathbb{P}(v \in A_M \setminus A_\varepsilon) \leq \frac{2\varepsilon}{\sigma}. \tag{2}$$

By a simple potential argument, there are indeed at most $2/\varepsilon$ tiles in $\mathcal{G}_\varepsilon$ that have non-trivial intersection with A_M (i.e., they intersect with A_M but are not fully contained in it). Moreover, each such tile has probability at most ε^2/σ by smoothness, and their union is a superset of $A_M \setminus A_\varepsilon$, so that $\mathbb{P}(v \in A_M \setminus A_\varepsilon) \leq 2\varepsilon/\sigma$; see Figure 1. The argument for A_ε is more subtle, as the payments induced by the grid mechanism may be considerably different than those computed by M, see again Figure 1. Since the profit is composed of the price paid by the buyer, corresponding to the vertical projection, and that paid to the seller, horizontal projection, we partition $[0,1]^2$ into columns C and rows R. A column is simply a rectangle of the type $(k\varepsilon,(k+1)\varepsilon] \times [0,1]$, while a row has the form $[0,1]\times[j\varepsilon,(j+1)\varepsilon)$. We can associate to each column (respectively row), an upper bound on the gap between the price paid by the buyer (respectively to the seller) in the two mechanisms: $\Delta_c = \sup_{v\in c\cap A_\varepsilon}\{q_\varepsilon(v) - q_M(v)\}$ and $\Delta_r = \sup_{v\in r\cap A_\varepsilon}\{p_M(v) - p_\varepsilon(v)\}$. We then have the following chain of inequalities, using p_ε and q_ε in place of $p_{\phi_\varepsilon(M)}$ and $q_{\phi_\varepsilon(M)}$:

$$\begin{aligned}
&\mathbb{E}\left[|\mathsf{profit}(M,v) - \mathsf{profit}(\phi_\varepsilon(M),v)|\mathbb{1}_{\{v\in A_\varepsilon\}}\right] \\
&\leq \mathbb{E}\left[\left(|q_\varepsilon(v) - q_M(v)| + |p_M(v) - p_\varepsilon(v)|\right)\mathbb{1}_{\{v\in A_\varepsilon\}}\right] \\
&\leq \mathbb{E}\left[\sum_{c\in C}|q_\varepsilon(v) - q_M(v)|\mathbb{1}_{\{v\in A_\varepsilon\cap c\}} + \sum_{r\in R}|p_M(v) - p_\varepsilon(v)|\mathbb{1}_{\{v\in A_\varepsilon\cap r\}}\right] \\
&\leq \sum_{c\in C}\Delta_c\mathbb{P}(v\in A_\varepsilon\cap c) + \sum_{r\in R}\Delta_r\mathbb{P}(v\in A_\varepsilon\cap r) && \text{(By def. of } \Delta_r \text{ and } \Delta_c\text{)} \\
&\leq \frac{\varepsilon}{\sigma}\left(\sum_{c\in C}\Delta_c + \sum_{r\in R}\Delta_r\right) \leq 4\frac{\varepsilon}{\sigma}, && \text{(By smoothness and } \mathcal{L}(c) = \mathcal{L}(r) = \varepsilon\text{)}
\end{aligned}$$

where the last inequality follows by arguing that both $\sum_c \Delta_c$ and $\sum_r \Delta_r$ are upper bounded by 2. The formal proof is reported in Claim 1. □

Claim 1. *It holds that* $\sum_{c\in C}\Delta_c \leq 2$ *and* $\sum_{r\in R}\Delta_r \leq 2$.

Proof. We prove the statement for the columns, the argument for the rows is analogous. Any column c can be written as $I_c \times [0,1]$, where $I_c = (k\varepsilon,(k+1)\varepsilon]$, for an integer k. Provided trade happens (i.e. $v \in A_\varepsilon \subseteq A_M$), the payment function q_M is constant in the buyer valuation v_b given the seller's v_s, and we focus without loss of generality on $v_b = 1$. By construction of the approximating mechanism, $q_\varepsilon(\cdot,1)$ is constant in I_c as a function of the seller valuation, and is equal to

$$Q_c = \varepsilon\left\lceil\frac{q_M((k+1)\varepsilon,1)}{\varepsilon}\right\rceil.$$

Notice that Q_c is the height of the lower side of the lowest tile fully contained in both c and A_M; if the intersection of the two does not contain any tile, then $Q_c = 1$. At this point, consider Δ_c. Since q_M is weakly increasing as a function of v_s (a consequence of the monotonicity of A_M), the gap between $q_\varepsilon(v_s,1)$ and $q_M(v_s,1)$ over the interval $(k\varepsilon,(k+1)\varepsilon]$ is maximized approaching the infimum of the interval. Therefore, taking the right-sided limit, we get $\Delta_c = Q_c - \lim_{v_s\to k\varepsilon^+} q_M(v_s,1)$.

Assuming it exists, i.e. $k \neq 0$, let us now consider the previous column, denoted as c_-, corresponding to the interval $I_{c_-} = ((k-1)\varepsilon, k\varepsilon]$. Repeating the same analysis, we get $Q_{c_-} = \varepsilon\left\lceil\frac{q_M(k\varepsilon,1)}{\varepsilon}\right\rceil$. By definition of the ceiling function, we have $Q_{c_-} - \varepsilon \leq q_M(k\varepsilon,1)$. Furthermore, by the monotonicity of q_M, its value at the boundary is bounded by its right limit: $q_M(k\varepsilon,1) \leq \lim_{v_s\to k\varepsilon^+} q_M(v_s,1)$. Together with the value of Δ_c, the above proves:

$$\Delta_c = Q_c - \lim_{v_s\to k\varepsilon^+} q_M(v_s,1) \leq Q_c - q_M(k\varepsilon,1) \leq Q_c - (Q_{c_-} - \varepsilon),$$

which implies $Q_c - Q_{c_-} \geq \Delta_c - \varepsilon$. If $k = 0$ and c_- does not exist, then $q_\varepsilon(0,1)$ is null by definition, with $Q_c \geq \Delta_c$ being trivially satisfied. To conclude, since $q_\varepsilon(\cdot,1)$ is monotone and its image is a subset of $[0,1]$, its total variation is bounded by 1. We then get, with $Q_{c_-} = 0$ when $I_c = (0,\varepsilon]$:

HIER-MECH

Environment: Repeated Bilateral Trade against a σ-smooth adversary on time horizon T
Construct the mechanism tree $\mathcal{T}$ with target precision $H = 1/2\lceil \log_2 T \rceil$
for each level $h < H$ and mechanism $M \in \mathcal{M}_h$ **do**
  initialize a local version of HEDGE $\mathcal{H}^h(M)$, with actions $\mathcal{N}_M$ and learning rate $\eta_h = 2^h/\sqrt{T}$
for time $t = 1, 2, \ldots, T$ **do**
  for each level $h < H$ and mechanism $M \in \mathcal{M}_h$ **do**
    Receive the prediction $\mathcal{H}^h_t(M)$ from the local version of HEDGE $\mathcal{H}^h(M)$
  Propose mechanism $M^t = \theta_t(M_\emptyset)$ and observe valuations (v^t_s, v^t_b)
  for each level $h < H$ and mechanism $M \in \mathcal{M}_h$ **do**
    Feed to $\mathcal{H}^h(M)$ rewards $\Delta^h_t(M') = 1/2 \cdot (\mathsf{profit}_t(\theta_t(M')) - \mathsf{profit}_t(M)) \, \forall \, M' \in \mathcal{N}_M$

$$1 \geq \sum_{c \in C} (Q_c - Q_{c_-}) \geq \sum_{c \in C} (\Delta_c - \varepsilon),$$

which implies, using $|C| \leq 1/\varepsilon$: $\sum_{c \in C} \Delta_c \leq 2$. □

4 Hierarchical Mechanism

In this section, we present our learning algorithm HIER-MECH and analyze its regret property. It instantiates multiple copies of the classical HEDGE algorithm [Freund and Schapire, 1997] as subroutines; so we point to Appendix D for a quick recap of the algorithm and of its guarantees.

Mechanism Tree. We introduce the mechanism tree $\mathcal{T}$, whose nodes are the 2^{-h}-grid mechanisms, for $h = 0, \ldots, H$ (see also Figure 3). For simplicity, we denote with $\mathcal{M}_h$ the class of 2^{-h}-grid mechanisms, and with $\phi_h(\cdot)$ (instead of $\phi_{2^{-h}}(\cdot)$) the generic 2^{-h} approximating mechanism in $\mathcal{M}_h$. The h-th level of this tree is given by the nodes in $\mathcal{M}_h$ (for simplicity, we ignore at all levels the suboptimal mechanism which always allocates, so that $\mathcal{M}_0$ only contains the mechanism $M_\emptyset$ which almost surely never allocates). We specify the edges of $\mathcal{T}$: there only exist edges going from one level to the next one, connecting the generic mechanism $M \in \mathcal{M}_h$ with the unique $\phi_{h-1}(M) \in \mathcal{M}_{h-1}$. For any node/mechanism M, we denote with $\mathcal{N}_M$ its children. We say that a mechanism is an internal node if it is not a leaf. Note that by construction, for each node $M \in \mathcal{M}_h$ there exists a unique root-to-node path in $\mathcal{T}$, namely, $M_\emptyset = \phi_0(M), \phi_1(M), \ldots, \phi_h(M) = M$.

The algorithm: HIER-MECH. Our algorithm, HIER-MECH, maintains a separate instantiation $\mathcal{H}^h(M)$ of the HEDGE algorithm for each internal node $M \in \mathcal{M}_h$ in the tree, and uses them to decide which mechanism to propose at each time step t. We refer to the pseudocode for further details. For each local instantiation of HEDGE to be well defined, we need to specify three things: the (a) learning rate, (b) the action space, and (c) the corresponding rewards. (a) As learning rate we use a level-dependent value $\eta_h = 2^h/\sqrt{T}$, while (b) the actions of the version of HEDGE $\mathcal{H}^h(M)$ supported on mechanism $M \in \mathcal{M}_h$ are the children $\mathcal{N}_M \subseteq \mathcal{M}_{h+1}$ of M in the mechanism tree $\mathcal{T}$. Finally, (c) the rewards are defined recursively starting from the leaves and moving up: for each child mechanism $M' \in \mathcal{N}_M$, the reward of playing M' is defined as $\Delta^h_t(M') = 1/2 \cdot (\mathsf{profit}_t(\theta_t(M')) - \mathsf{profit}_t(M))$, where the target function $\theta_t(\cdot)$ is also defined recursively ($\mathcal{H}^h_t(M)$ denotes the random action played by the local hedge $\mathcal{H}^h(M)$ at time t):

$$\begin{cases} \theta_t(M) = & M \quad \text{for a leaf} \\ \theta_t(M) = & \theta_t(\mathcal{H}^h_t(M)) \quad \text{otherwise} \end{cases}$$

Stated differently, the reward of choosing mechanism M' in the local HEDGE version in M is given by the *residual* profit achievable by following the choices of the local HEDGE instantiations down the subtree rooted in M, up to the leaf $\theta_t(M')$. Given all these local HEDGE, the learner actually proposes to the agents the mechanism $M^t = \theta_t(M_\emptyset)$. Namely, at each time step, it follows the root-to-leaf path sampled by the local instantiations of HEDGE, and chooses the resulting leaf mechanism. Once (v^t_s, v^t_b) is revealed, then HIER-MECH can update the internal state of each local HEDGE.

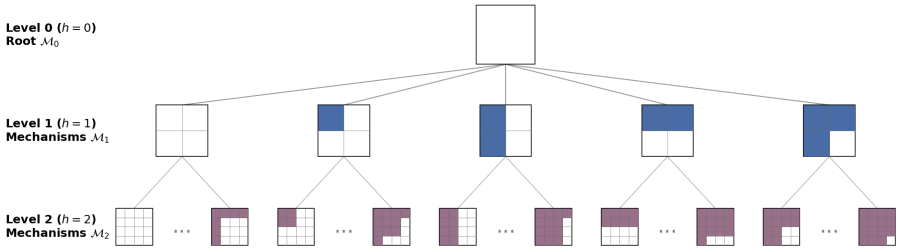

Figure 3: The first levels of the mechanism tree. We are only showing a subset of level $h = 2$.

The analysis of HIER-MECH. We denote with M^t the random mechanism played by the algorithm at the generic time t, and with $M^\star$ the best mechanism on the instance*. We use $\phi_h^\star$ to denote the generic approximating mechanism $\phi_h(M^\star)$, and with $\mathcal{H}^h$ (instead of $\mathcal{H}^h(\phi_h^\star)$) the corresponding version of HEDGE. In the analysis of HIER-MECH we only need to consider these local versions of HEDGE, one for each level of the tree, whose properties are summarized in the following lemma.

Lemma 3 (Local Regret)**.** *For each level $h < H$, we have the following bound on the regret of the version of* HEDGE $\mathcal{H}^h$ *at that level, where expectations are taken with respect to the algorithm randomization and the distribution of the valuations at each time step:*

$$\sum_{t=1}^{T} \mathbb{E}\left[\Delta_t^h(\phi_{h+1}^\star) - \Delta_t^h(\mathcal{H}_t^h)\right] \leq \frac{7}{\sigma}\sqrt{T}.$$

Proof of Lemma 3. Consider the generic level $h < H$, the HEDGE algorithm $\mathcal{H}^h$ corresponds to node $\phi_h^\star$, and one of its actions is the next element $\phi_{h+1}^\star$ in the root-to-leaf path to the best mechanism $M^\star$. For brevity, denote with $\mathcal{N}_{h+1}$ the children of $\phi_h^\star$.

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{E}\left[\Delta_t^h(\phi_{h+1}^\star) - \Delta_t^h(\mathcal{H}_t^h)\right] &\leq \max_{M \in \mathcal{N}_{h+1}} \sum_{t=1}^{T} \mathbb{E}\left[\Delta_t^h(M) - \Delta_t^h(\mathcal{H}_t^h)\right] && \text{(As } \phi_{h+1}^\star \in \mathcal{N}_{h+1}\text{)} \\
&\leq \frac{\log|\mathcal{N}_{h+1}|}{\eta_h} + \eta_h \cdot \sum_{t=1}^{T} \mathbb{E}\left[\Delta_t^h(\mathcal{H}_t^h)^2\right] && \text{(HEDGE guarantees)} \\
&\leq 4\frac{2^h}{\eta_h} + \eta_h \sum_{t=1}^{T} \mathbb{E}\left[|\Delta_t^h(\mathcal{H}_t^h)|\right] && \text{(Lemma 1, } |\Delta_t^h(M)| \leq 1\text{)}
\end{aligned}$$

Note, the second inequality follows by the standard analysis of the HEDGE algorithm; for completeness, we report a formal proof in Appendix D. Now, recall the definition of $\Delta_t^h(M') = {}^1\!/_2 \cdot (\mathsf{profit}_t(\theta_t(M')) - \mathsf{profit}_t(\phi_h^\star))$ for any $M' \in \mathcal{N}_{h+1}$. By construction, $\theta_t(M')$ is a leaf in the sub-tree rooted in $\phi_h^\star$, therefore $\phi_h(\theta_t(M')) = \phi_h^\star$. By Lemma 2, it then follows that the expected absolute profit difference between any fixed leaf of this subtree and $\phi_h^\star$ is at most ${}^6\!/_\sigma \cdot 2^{-h}$.

Conditioning on the history and all random choices made before observing the valuations at time t then fixes $\theta_t(\mathcal{H}_t^h)$, while leaving the valuation distribution unchanged, by independence of the distribution at time t from the previous ones. Applying Lemma 2 conditionally and then averaging thus yields $\mathbb{E}\left[|\Delta_t^h(\mathcal{H}_t^h)|\right] \leq {}^3\!/_\sigma \cdot 2^{-h}$. Plugging in this inequality in the previous display yields:

$$\sum_{t=1}^{T} \mathbb{E}\left[\Delta_t^h(\phi_{h+1}^\star) - \Delta_t^h(\mathcal{H}_t^h)\right] \leq 4 \cdot \frac{2^h}{\eta_h} + 3 \cdot \frac{\eta_h}{\sigma} 2^{-h} \cdot T \leq \frac{7}{\sigma}\sqrt{T},$$

where the last inequality follows by our choice of the level-dependent learning rate. □

The lemma ensures that the local regret suffered along each edge of the root-to-leaf path to $\phi_H^\star$ enjoys the same rate! Our choice of the learning rate finds the right trade-off between the number of available

*Although it is possible to prove with some work that such max is indeed attained, one can alternatively define $M^\star$ as any mechanism whose expected profit is optimal up to some small constant term, with this not changing the analysis.

mechanisms (higher h corresponds to larger number of children) and the expected magnitude of the residual profit (higher h corresponds to lower magnitude, by Lemma 2). We now relate the actual profit of HIER-MECH with the benchmark $M^\star$ (more precisely to its approximation $\phi_H^\star$). This is the core of the argument, and is analogous to what happens in probabilistic chaining.

Lemma 4 (Chaining the Profit)**.** *For each level $0 < h \leq H$, we have the following regret bound:* $\sum_{t=1}^{T} \mathbb{E}\left[\mathsf{profit}_t(\theta_t(\phi_h^\star)) - \mathsf{profit}_t(M^t)\right] \leq \frac{14}{\sigma} h\sqrt{T}$.

Proof. We prove the result by induction on the level $h \leq H$. From Lemma 3 applied at the version of HEDGE corresponding to the root, we have: $\sum_{t=1}^{T} \mathbb{E}\left[\Delta_t^0(\phi_1^\star) - \Delta_t^0(\mathcal{H}_t^0)\right] \leq \frac{7}{\sigma}\sqrt{T}$. For any $M \in \mathcal{N}_{M_\emptyset}$, $\Delta_t^0(M) = 1/2 \cdot \mathsf{profit}_t(\theta_t(M))$ almost surely, as the profit of the root mechanism $M_\emptyset$ is almost surely 0. Moreover, by design of the algorithm, it holds that $2\Delta_t^0(\mathcal{H}_t^0) = \mathsf{profit}_t(M^t)$, while $2\Delta_t^0(\phi_1^\star) = \mathsf{profit}_t(\theta_t(\phi_1^\star))$, since $\phi_1^\star \in \mathcal{N}_{M_\emptyset}$. Combining these facts yields the base case:

$$\sum_{t=1}^{T} \mathbb{E}\left[\mathsf{profit}_t(\theta_t(\phi_1^\star)) - \mathsf{profit}_t(M^t)\right] \leq \frac{14}{\sigma}\sqrt{T}.$$

Assume now by inductive hypothesis that the statement holds for a generic level $h < H$, it means that, recalling that $\mathcal{H}^h$ refers to the HEDGE instance for node $\phi_h^\star$:

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{E}\left[\mathsf{profit}_t(M^t)\right] &\geq -\frac{14}{\sigma} h\sqrt{T} + \sum_{t=1}^{T} \mathbb{E}\left[\mathsf{profit}_t(\theta_t(\phi_h^\star)) \pm \mathsf{profit}_t(\theta_t(\phi_{h+1}^\star)) \pm \mathsf{profit}_t(\phi_h^\star)\right] \\
&= -\frac{14}{\sigma} h\sqrt{T} + \sum_{t=1}^{T} \mathbb{E}\left[\mathsf{profit}_t(\theta_t(\phi_{h+1}^\star))\right] + 2\sum_{t=1}^{T} \mathbb{E}\left[\Delta_t^h(\mathcal{H}_t^h) - \Delta_t^h(\phi_{h+1}^\star)\right] \\
&\geq -\frac{14}{\sigma}(h+1)\sqrt{T} + \sum_{t=1}^{T} \mathbb{E}\left[\mathsf{profit}_t(\theta_t(\phi_{h+1}^\star))\right],
\end{aligned}$$

where the equality follows by the definition of Δ_t^h, and the last inequality by Lemma 3. □

Theorem 1. *Consider the problem of profit maximization in repeated bilateral trade against a σ-smooth adversary, with σ possibly unknown. We then have the following regret bound:* $R_T(\text{HIER-MECH}) \leq \frac{20}{\sigma}\sqrt{T}\log T$.

Proof. Recall that $\phi_H^\star$ is the approximating mechanism with precision 2^{-H} of the optimal mechanism $M^\star$. We have the following inequalities, which conclude the proof as $H = 1/2\lceil \log_2 T \rceil$:

$$\begin{aligned}
\frac{14}{\sigma} H\sqrt{T} &\geq \sum_{t=1}^{T} \mathbb{E}\left[\mathsf{profit}_t(\theta_t(\phi_H^\star)) - \mathsf{profit}_t(M^t)\right] && \text{(By Lemma 4 on the last level } H\text{)} \\
&= \sum_{t=1}^{T} \mathbb{E}\left[\mathsf{profit}_t(\phi_H^\star) - \mathsf{profit}_t(M^t) \pm \mathsf{profit}_t(M^\star)\right] && \text{(As } \phi_H^\star \text{ is a leaf)} \\
&\geq \sum_{t=1}^{T} \mathbb{E}\left[\mathsf{profit}_t(M^\star) - \mathsf{profit}_t(M^t)\right] - \frac{6}{\sigma} \cdot 2^{-H} T && \text{(By Lemma 2)} \\
&\geq \sum_{t=1}^{T} \mathbb{E}\left[\mathsf{profit}_t(M^\star) - \mathsf{profit}_t(M^t)\right] - \frac{6}{\sigma} \cdot \sqrt{T}. && \square
\end{aligned}$$

5 A Reduction from Joint Ads to Bilateral Trade

In this section, we provide a high-level intuition behind our reduction from joint ads to profit maximization in bilateral trade. We refer to Appendix C.2 for a definition of the joint ads model, while the formal proof of the reduction is in Appendix F.

Joint ads. Here, two buyers cooperate in an auction for a non-excludable good: either they both get it, or none does. The mechanism designer wants to maximize the profit/revenue extracted from

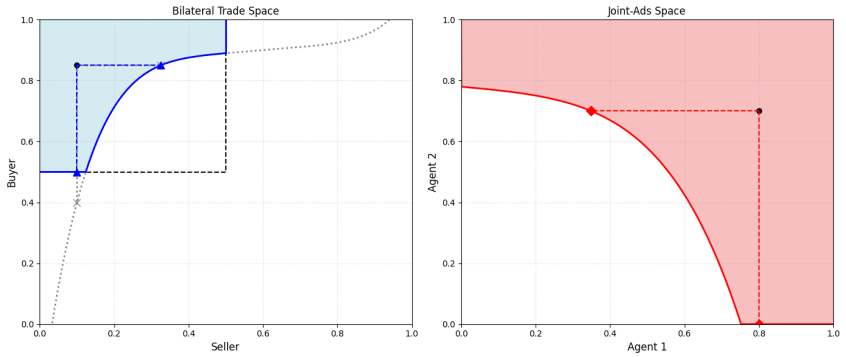

Figure 4: A visualization of our reduction from the joint ads problem.

the trade, while enforcing DSIC and IR constraints. The main difference from bilateral trade is the "direction of monotonicity": allocation regions need to be "north-east" closed, while payments are given by the "south" and "west" projections[†]. For a visualization, we refer to the second picture of Figure 4. The learning protocol is analogous to bilateral trade. The agents are characterized by private valuations (v_1^t, v_2^t), drawn from a σ-smooth distribution. The learner proposes mechanism M^t with allocation A^t and two payments p_1^t and p_2^t, the trade happens if $(v_1^t, v_2^t) \in A^t$ and the revenue is: $\mathsf{rev}_t(M^t) = p_1^t(v_1^t, v_2^t) + p_2^t(v_1^t, v_2^t)$. While the two problems share some similarities, in bilateral trade we may have negative profit for some trades, while this cannot happen for joint ads. In both problems, the trade not happening results in 0 profit/revenue, thus introducing an asymmetry.

Our Reduction. To overcome this issue, we restrict the bilateral trade instance to $[0, 1/2] \times [1/2, 1]$, so that no trade results in *negative* profit (as the buyer pays at least $1/2$, and the seller receives at most $1/2$). Luckily, by monotonicity this restriction goes "in the right" direction for the reduction to happen. Given a learning algorithm $\mathcal{A}_{\text{BT}}$ for bilateral trade, we now show how to derive $\mathcal{A}_{\text{JA}}$ for joint ads: at time t, we query mechanism M_{BT}^t from $\mathcal{A}_{\text{BT}}$, we then *clip* it to the square of interest $[0, 1/2] \times [1/2, 1]$ (see first plot in Figure 4), and apply the inverse of the following map $f : (x, y) \to (1/2 \cdot (1 - x), 1/2 \cdot (1 + y))$. The resulting allocation region A_{JA}^t induces the mechanism M_{JA}^t. The joint ads learner proposes M_{JA}^t and observes the agents' valuations (v_1^t, v_2^t), which are then mapped back to bilateral trade via f. As corollaries, we prove in Appendix F the following:

Corollary 1 (A lower bound for bilateral trade)**.** *Consider the problem of profit maximization in repeated bilateral trade against a σ-smooth adversary, for $\sigma \in (0, 1/48)$. Then, for any learning algorithm $\mathcal{A}$, the following lower bound holds: $R_T(\mathcal{A}) \geq \frac{3}{128}\sqrt{T}$.*

Corollary 2 (An upper bound for joint ads)**.** *Consider the problem of revenue maximization in repeated joint ads against a σ-smooth adversary, with σ possibly unknown. There exists a learning algorithm which suffers regret at most $\frac{160}{\sigma}\sqrt{T}\log T$.*

6 Conclusion and Future Directions

In this paper, we characterize the minimax regret rates for profit maximization in bilateral trade for the σ-smooth adversary, up to lower order terms. We achieve this result via an algorithmic chaining approach, for a suitable L^1-net construction on the mechanism space; we envision that this and similar approaches have a vast potential in economically motivated learning problems, and beyond. We further formalize a reduction from the joint ads problem, a related mechanism design task. This allows us to provide the same tight results also for this new problem in the smooth adversarial setting. We leave open two natural and exciting research questions that stem from our results: understanding the right dependency on the smoothness parameter σ in the regret rate, and assessing whether the same rates can be attained in time polynomial in T, as opposed to polynomial in the number of experts maintained.

[†]Recall: in bilateral trade, the monotonicity is in the north-west direction.

Acknowledgments

The work of SDG and FF was supported in part by the MUR PRIN grant 2022EKNE5K (Learning in Markets and Society), while CS was supported by a Google Research Award. SDG was also supported in part by the Institute for Complex Systems (Italian National Research Council).

References

Gagan Aggarwal, Ashwinkumar Badanidiyuru, Paul Dütting, and Federico Fusco. Selling joint ads: A regret minimization perspective. In *EC*, pages 164–194. ACM, 2024.

Martin Anthony and Peter L. Bartlett. *Neural Network Learning: Theoretical Foundations*. Cambridge University Press, 1999.

Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory Comput.*, 8(6):121–164, 2012.

Yossi Azar, Amos Fiat, and Federico Fusco. An α-regret analysis of adversarial bilateral trade. *Artif. Intell.*, 337:104231, 2024.

Francesco Bacchiocchi, Matteo Bollini, Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. Online bayesian persuasion without a clue. In *NeurIPS*, 2024.

François Bachoc, Tommaso Cesari, and Roberto Colomboni. A parametric contextual online learning theory of brokerage. In *ICML*, volume 267 of *Proceedings of Machine Learning Research*, pages 2293–2309. PMLR, 2025.

Shai Ben-David, Dávid Pál, and Shai Shalev-Shwartz. Agnostic online learning. In *COLT*, 2009.

Martino Bernasconi, Matteo Castiglioni, Andrea Celli, Alberto Marchesi, Francesco Trovò, and Nicola Gatti. Optimal rates and efficient algorithms for online bayesian persuasion. In *ICML*, volume 202 of *Proceedings of Machine Learning Research*, pages 2164–2183. PMLR, 2023.

Martino Bernasconi, Matteo Castiglioni, Andrea Celli, and Federico Fusco. No-regret learning in bilateral trade via global budget balance. In *STOC*, pages 247–258. ACM, 2024.

Moïse Blanchard. Agnostic smoothed online learning. In *STOC*, pages 1997–2006. ACM, 2025.

Adam Block and Max Simchowitz. Efficient and near-optimal smoothed online learning for generalized linear functions. In *NeurIPS*, 2022.

Adam Block, Yuval Dagan, and Alexander Rakhlin. Majorizing measures, sequential complexities, and online learning. In *COLT*, volume 134 of *Proceedings of Machine Learning Research*, pages 587–590. PMLR, 2021.

Adam Block, Yuval Dagan, Noah Golowich, and Alexander Rakhlin. Smoothed online learning is as easy as statistical learning. In *COLT*, volume 178 of *Proceedings of Machine Learning Research*, pages 1716–1786. PMLR, 2022.

Adam Block, Alexander Rakhlin, and Abhishek Shetty. On the performance of empirical risk minimization with smoothed data. In *COLT*, volume 247 of *Proceedings of Machine Learning Research*, pages 596–629. PMLR, 2024.

Avrim Blum, Vijay Kumar, Atri Rudra, and Felix Wu. Online learning in online auctions. In *SODA*, pages 202–204. SIAM, 2003.

Nicolò Cesa-Bianchi, Claudio Gentile, and Yishay Mansour. Regret minimization for reserve prices in second-price auctions. *IEEE Trans. Inf. Theory*, 61(1):549–564, 2015.

Nicolò Cesa-Bianchi, Pierre Gaillard, Claudio Gentile, and Sébastien Gerchinovitz. Algorithmic chaining and the role of partial feedback in online nonparametric learning. In *COLT*, volume 65 of *Proceedings of Machine Learning Research*, pages 465–481. PMLR, 2017.

Nicolò Cesa-Bianchi, Tommaso Cesari, Roberto Colomboni, Federico Fusco, and Stefano Leonardi. The role of transparency in repeated first-price auctions with unknown valuations. In *STOC*, pages 225–236. ACM, 2024a.

Nicolò Cesa-Bianchi, Tommaso Cesari, Roberto Colomboni, Federico Fusco, and Stefano Leonardi. Regret analysis of bilateral trade with a smoothed adversary. *J. Mach. Learn. Res.*, 25:234:1–234:36, 2024b.

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, 2024c.

Nicolò Cesa-Bianchi, Tommaso Cesari, Roberto Colomboni, Luigi Foscari, and Vinayak Pathak. Market making without regret. In *COLT*, volume 291 of *Proceedings of Machine Learning Research*, pages 799–837. PMLR, 2025.

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.

Paul Duetting, Guru Guruganesh, Jon Schneider, and Joshua Ruizhi Wang. Optimal no-regret learning for one-sided lipschitz functions. In *ICML*, volume 202 of *Proceedings of Machine Learning Research*, pages 8836–8850. PMLR, 2023.

Naveen Durvasula, Nika Haghtalab, and Manolis Zampetakis. Smoothed analysis of online non-parametric auctions. In *EC*, pages 540–560. ACM, 2023.

Zhe Feng, Chara Podimata, and Vasilis Syrgkanis. Learning to bid without knowing your value. In *EC*, pages 505–522. ACM, 2018.

Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.*, 55(1):119–139, 1997.

Wenshuo Guo, Michael I. Jordan, and Ellen Vitercik. No-regret learning in partially-informed auctions. In *ICML*, volume 162 of *Proceedings of Machine Learning Research*, pages 8039–8055. PMLR, 2022.

Nika Haghtalab, Tim Roughgarden, and Abhishek Shetty. Smoothed analysis of online and differentially private learning. In *NeurIPS*, 2020.

Nika Haghtalab, Yanjun Han, Abhishek Shetty, and Kunhe Yang. Oracle-efficient online learning for smoothed adversaries. In *NeurIPS*, 2022.

Nika Haghtalab, Tim Roughgarden, and Abhishek Shetty. Smoothed analysis with adaptive adversaries. *J. ACM*, 71(3):19, 2024.

Robert Kleinberg, Aleksandrs Slivkins, and Eli Upfal. Bandits and experts in metric spaces. *J. ACM*, 66(4):30:1–30:77, 2019.

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, 2003.

Rachitesh Kumar, Jon Schneider, and Balasubramanian Sivan. Strategically-robust learning algorithms for bidding in first-price auctions. In *EC*, page 893. ACM, 2024.

Anna Lunghi, Matteo Castiglioni, and Alberto Marchesi. Better regret rates in bilateral trade via sublinear budget violation. In *SODA*, pages 6494–6536. SIAM, 2026.

Roger B. Myerson. Optimal auction design. *Math. Oper. Res.*, 6(1):58–73, 1981.

Roger B. Myerson and Mark A. Satterthwaite. Efficient mechanisms for bilateral trading. *J. Econ. Theory*, 29(2):265–281, 1983.

David Pollard. *Convergence of Stochastic Processes*. Springer Series in Statistics. Springer-Verlag, 1984.

Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Online learning: Stochastic, constrained, and smoothed adversaries. In *NIPS*, pages 1764–1772, 2011.

Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Online learning via sequential complexities. *J. Mach. Learn. Res.*, 16:155–186, 2015.

Vinod Raman, Unique Subedi, and Ambuj Tewari. Smoothed online classification can be harder than batch classification. In *NeurIPS*, 2024.

Tim Roughgarden, editor. *Beyond the Worst-Case Analysis of Algorithms*. Cambridge University Press, 2020.

Tim Roughgarden and Joshua R. Wang. Minimizing regret with multiple reserves. *ACM Trans. Econ. Comput.*, 7(3):17:1–17:18, 2019.

Michel Talagrand. *Upper and lower bounds for stochastic processes*, volume 60 of *A Series of Modern Surveys in Mathematics*. Springer Nature, 2nd edition, 2021.

Vladimir N. Vapnik and Alexey Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab. Appl.*, 16(2):264–280, 1971.

Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2nd edition, 2026.

William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *J. Finance*, 16(1): 8–37, 1961.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes].

   Justification: All claims are supported by formal theorems proved in the paper.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes].

   Justification: We clearly state our mathematical assumptions and the limitations of our work.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes].

Justification: Yes, the paper contains mathematical proofs of all our claimed results.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A].

   Justification: the paper does not include experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [N/A].

Justification: the paper does not include experiments.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A].

   Justification: the paper does not include experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A].

   Justification: the paper does not include experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A].

   Justification: the paper does not include experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes].

   Justification: we have reviewed the NeurIPS Code of Ethics and our paper conforms, in every respect, to it.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [N/A].

    Justification: The paper advances the theoretical machine learning research and we do not see any specific societal impact.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A].

    Justification: the paper poses no such risks.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [N/A].

    Justification: the paper does not use existing assets.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A].

    Justification: the paper does not release new assets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A].

    Justification: the paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A].

    Justification: the paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A].

Justification: the core method development in this research does not involve LLMs as any important, original, or non-standard components.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.

A Additional Related Work

Repeated Bilateral Trade. The bilateral trade problem has been extensively investigated in the regret minimization setting when the underlying goal is to maximize welfare (or, equivalently, gain from trade). A list of work [Cesa-Bianchi et al., 2024c,b, Azar et al., 2024, Bernasconi et al., 2024, Lunghi et al., 2026] has characterized the minimax regret rates achievable under various feedback models, data generation models and budget balance constraints (when the goal is welfare, it makes sense to impose the additional constraint for the broker *not to lose money*). Note, in all this literature, the regret benchmark is the best fixed price, as opposed to our rich class of mechanisms. Conversely, in profit maximization we need access to the agents' valuations to compute payments, while for fixed-price mechanisms it is natural to study less informative feedback structures. We observe that smoothness has already played an important role in these works, as an important tool to avoid pathological "needle-in-a-haystack" lower bounds while remaining a meaningful and challenging problem [Cesa-Bianchi et al., 2024c,b].

Smoothed Online Learning. In the last ten years, a growing body of work — starting from Rakhlin et al. [2011] — has been focusing on the smoothness assumption to overcome the common impossibility results from adversarial online learning. Indeed, while general online learnability is characterized via very strong sequential complexity measures like the Littlestone dimension [Ben-David et al., 2009, Rakhlin et al., 2015, Block et al., 2021], the smoothness assumption for the adaptive adversary collapses the statistical hardness of online classification and regression back to the well-known VC/fat shattering dimension [Haghtalab et al., 2020, 2024, Block et al., 2022]. Notably, this is not true for multiclass classification with unbounded label space, when facing a slightly stronger smooth adversary [Raman et al., 2024]. More recently, Blanchard [2025] and Block et al. [2024] enriched and generalized this information-theoretical viewpoint by investigating the smooth online learning setting when the underlying measure is not assumed to be the Lebesgue measure and it is not assumed to be known to the learner, again obtaining VC-dependent regret rates.

Since our hypothesis class is not arbitrary and our learning objective (i.e. the optimal mechanism) is dependent on the valuation distribution, our setting is not captured by the lower bounds proven in this line of work; this is part of the reason why learning is for us achievable even if the fat shattering dimension is unbounded. Moreover, since the fat shattering dimension is unbounded (see Appendix B), the positive results and associated techniques are mostly inapplicable to our case. However, we do encounter the same information-theoretic separation between smooth and general adversaries, in the sense of smoothness being the enabler of a non-trivial minimax rate.

From a computational perspective, a different but related line of work considers the problem of attaining optimal statistical rates under the smooth assumption while minimizing the time complexity of the algorithm used to attain them, either explicitly — for certain model specifications [Block and Simchowitz, 2022] — or implicitly in terms of calls to an empirical risk minimization oracle [Block et al., 2022, Haghtalab et al., 2022].

Mechanism Design and Online Learning. Modern computational economics and learning theory have a strong intersection, mostly starting from Kleinberg and Leighton [2003], with a variety of economic problems which have been studied through this lens. For example, auctions [Blum et al., 2003, Roughgarden and Wang, 2019, Kumar et al., 2024, Guo et al., 2022], market making [Cesa-Bianchi et al., 2025], contracts [Duetting et al., 2023], brokerage [Bachoc et al., 2025] and Bayesian persuasion [Bacchiocchi et al., 2024, Bernasconi et al., 2023] were all investigated from the online learning perspective.

B On the Impossibility of a Deterministic L^∞-net

In this appendix, we comment on the challenges arising in applying directly the analysis from Cesa-Bianchi et al. [2017] to our problem. In particular, to apply such results, we would need a finite family of mechanisms $\mathcal{M}_\varepsilon^\infty$ satisfying the following L^∞-net guarantee: for any mechanism $M \in \mathcal{M}$ there should exist an $M_\varepsilon \in \mathcal{M}_\varepsilon^\infty$ such that

$$|\mathsf{profit}(M, v) - \mathsf{profit}(M_\varepsilon, v)| \leq \varepsilon \quad \forall\, v \in [0,1]^2 \tag{3}$$

This net property is unattainable due to the strongly discontinuous nature of the profit function. For any M_ε *different than* M, the symmetric difference of the allocation regions contains at least one valuation v for which one mechanism extracts possibly non-null profit and the other one misses the trade.

We have two ways of overcoming this impossibility: either focusing on a still-deterministic but weaker notion of L^p bound, or by requiring a less stringent *in-expectation bound* using smoothness. The latter is what we derive in the main body (we restate below the relevant lemma for convenience):

Lemma 2. *Let v be a σ-smooth random variable on $[0,1]^2$. For every $\varepsilon > 0$ we have the following inequality:* $\mathbb{E}\left[|\mathsf{profit}(M, v) - \mathsf{profit}(\phi_\varepsilon(M), v)|\right] \leq 6 \cdot \varepsilon/\sigma$.

We devote the rest of the appendix to argue that the first solution proposed, namely a weaker deterministic L^p bound is hopeless (to some extent, this is reminiscent of the non-constructive nets used in Haghtalab et al. [2024]). Such notion would read as follows. For any fixed choice of n points S in $[0,1]^2$ there exists a sample-dependent $\mathcal{M}^p_\varepsilon(S)$ such that for every $M \in \mathcal{M}$ there exist an $M^p_\varepsilon \in \mathcal{M}^p_\varepsilon(S)$ verifying:

$$\left(\frac{1}{n}\sum_{v\in S}|\mathsf{profit}(M, v) - \mathsf{profit}(M^p_\varepsilon, v)|^p\right)^{1/p} \leq \varepsilon.$$

Note, this notion is weaker than Equation (3). Standard combinatorial arguments tell us that one can always find a finite $\mathcal{M}^p_\varepsilon(S)$ net respecting the above property, whose cardinality is, however, *exponential* in $|S|$. Such trivial bound does not imply any non-trivial learning guarantees.

Unfortunately, it turns out that it is impossible to get a sub-exponential net. Indeed, covering/packing duality (see Theorem 12.1 in Anthony and Bartlett [1999]) combined with the fact that the fat-shattering dimension is unbounded implies the non-existence of better-than-exponential L^p-nets.

We now prove that the fat-shattering dimension of the problem is unbounded. The fat shattering dimension [e.g. Chapter 11 in Anthony and Bartlett, 1999] is a "scale-sensitive" generalization of the VC dimension from Vapnik and Chervonenkis [1971] from binary classification. It is a crucial tool to *characterize the size of nets* for real-valued functions[‡], and it is much more general than the stricter pseudodimension [Pollard, 1984], which is unbounded (and thus useless) even for standard classes of functions (like the ones with bounded variation).

Definition 5 (γ-fat shattering)**.** *Let $\mathcal{F}$ be a class of functions mapping from $\mathcal{X}$ to $\mathbb{R}$, and let $\gamma > 0$. A set of points $S = \{x_1, x_2, \ldots, x_m\} \subset \mathcal{X}$ is said to be γ-fat shattered by $\mathcal{F}$ if there exists a witness vector $s = (s_1, s_2, \ldots, s_m) \in \mathbb{R}^m$ such that for every Boolean assignment $z \in \{-1, 1\}^m$, there exists a function $f \in \mathcal{F}$ satisfying, for all $i \in [m]$:*

$$\begin{aligned} f(x_i) &\geq s_i + \gamma \quad \textit{if } z_i = 1 \\ f(x_i) &\leq s_i - \gamma \quad \textit{if } z_i = -1 \end{aligned}$$

The γ-fat shattering dimension of $\mathcal{F}$, $\mathrm{fat}_\mathcal{F}(\gamma)$ *is the cardinality of the largest $A \subseteq \mathcal{X}$ which can be γ-fat shattered. If no upper bound can be found, the fat shattering dimension is defined to be infinite.*

Proposition 2 (Unbounded fat shattering)**.** *For any $\gamma \leq 1/4$, the γ-fat shattering dimension of the profit functions in bilateral trade $\mathcal{F}_\mathcal{M}$ is unbounded.*

Proof. Fix an arbitrary integer m and fix any set $S = \{v_i\}_{i=1}^m \subset C = \left\{(x, y) : x = y - \frac{1}{2}\right\} \cap (0,1)^2$. Any valuation vector in S lies on C, meaning that the difference between its entries is always $1/2$.

Let us fix an arbitrary vector of signs z: this induces a partition of S into S' and $S \setminus S'$. We can build a mechanism M' and a monotone allocation region $A_{M'}$ such that the corresponding profit function $\mathsf{profit}(M', \cdot) \in \mathcal{F}_\mathcal{M}$ assigns $1/2$ to points in S' and 0 otherwise. More precisely, let $(x_i, y_i) = v_i$ and $R_i = [0, x_i] \times [y_i, 1]$ for any $i \in [m]$; we define $A_{M'} = \cup_{i:\, v_i \in S'} R_i$.

If we take $s \in \mathbb{R}^m$ to be the constant vector with entries equal to $1/4$, we immediately have that $\mathsf{profit}(M', v_i) - s_i \geq 1/4$ if $z_i = 1$ and $\mathsf{profit}(M', v_i) - s_i \leq -1/4$ if $z_i = -1$. No assumption

[‡]An in-depth discussion about covering and packing numbers is out of scope for this paper, but we refer to Anthony and Bartlett [1999] for a good overview.

was made on z, so the same holds for any choice of z. Therefore, S is $1/4$-fat shattered by $\mathcal{F}$. The procedure is valid $\forall\, m \in \mathbb{N}$: $\text{fat}_{\mathcal{F}_{\mathcal{M}}}(1/4)$ is thus unbounded. The claim then follows by noting that $\text{fat}_{\mathcal{F}}(\gamma)$ is monotone decreasing in γ, for any function class $\mathcal{F}$. □

C Mechanism Design Background

In this appendix, we recall some of the basic notions from mechanism design about bilateral trade and the joint-ads problem. We focus particularly on the known geometric characterizations of DSIC and IR mechanisms utilized throughout our work.

C.1 Bilateral Trade

In the static bilateral trade problem, a seller and a buyer interact to exchange a single good. The seller's and buyer's (private) valuations for the good are $v_s \in [0,1]$ and $v_b \in [0,1]$, respectively. Seller and buyer submit not necessarily truthful bids $b_s \in [0,1]$ and $b_b \in [0,1]$ to a broker running mechanism M. A mechanism is characterized by an allocation region $A \subseteq [0,1]^2$, and pricing rules $p, q : [0,1]^2 \to [0,1]$.

The trade happens if and only if the bids (b_s, b_b) belong to the allocation region A, while the payments are made according to p and q. Payments are from the broker to the seller, and from the buyer to the broker: the seller thus receives $p(b_s, b_b)$, while the buyer pays $q(b_s, b_b)$.

The buyer's and seller's utility with valuations v_b and v_s are, when bidding $(b_b, b_s) \in [0,1]^2$:

$$U_s(b_s, b_b) = v_s - \mathbb{1}_{\{(b_s,b_b)\in A\}} \cdot v_s + p(b_s, b_b)$$
$$U_b(b_s, b_b) = \mathbb{1}_{\{(b_s,b_b)\in A\}} \cdot v_b - q(b_s, b_b)$$

We consider dominant-strategy incentive compatible (DSIC) and individually rational (IR) mechanisms. With this restriction, each player maximizes his utility with a truthful bid regardless of the other player's bid, and the utility from participating in the mechanism is at least as high as that from not participating in the mechanism. In formulas,

$$\begin{aligned}
\text{DSIC:}\quad & U_s(v_s, b_b) \geq U_s(b_s, b_b) && \forall\, v_s \in [0,1], (b_s, b_b) \in [0,1]^2 \\
& U_b(b_s, v_b) \geq U_b(b_s, b_b) && \forall\, v_b \in [0,1], (b_s, b_b) \in [0,1]^2 \\
\text{IR :}\quad & U_s(v_s, b_b) \geq v_s,\ U_b(b_s, v_b) \geq 0 && \forall\, (v_s, v_b) \in [0,1]^2, (b_s, b_b) \in [0,1]^2
\end{aligned}$$

As a result of the aforementioned requirements, we obtain Proposition 1. A comprehensive proof of the statement, utilizing standard tools from Myerson [1981], Myerson and Satterthwaite [1983], can be found in Di Gregorio et al. [2025].

Definition 1 (Monotone Allocation Region & Myerson Payments)**.** *An allocation region $A \subseteq [0,1]^2$ is monotone if for any $x = (x_1, x_2) \in A$ and $y = (y_1, y_2) \in [0,1]^2$ with $x_1 \geq y_1$ and $x_2 \leq y_2$, it holds that $y \in A$. A mechanism is monotone if its allocation region is monotone. Given a monotone allocation region, then the associated Myerson payments are defined as follows:*

$$\begin{aligned}
p(v_s, v_b) =& \mathbb{1}_{\{(v_s,v_b)\in A\}} \cdot \max\{x \in [0,1] \mid (x, v_b) \in A\} && \forall\, (v_s, v_b) \in [0,1]^2 \\
q(v_s, v_b) =& \mathbb{1}_{\{(v_s,v_b)\in A\}} \cdot \min\{y \in [0,1] \mid (v_s, y) \in A\} && \forall\, (v_s, v_b) \in [0,1]^2
\end{aligned}$$

Proposition 1. *A mechanism for bilateral trade is dominant-strategy incentive compatible and individually rational if and only if its allocation region is monotone and uses Myerson payments.*

As stated in the main body, we denote the family of mechanisms characterized by the above proposition as $\mathcal{M}$. Notice that the DSIC and IR conditions allow $p(v_s, v_b) \geq 0$ and $q(v_s, v_b) \leq 0$ when there is no trade[§], which is more general than Definition 1. Specifically, the constraints only define payments up to a constant offset. However, considering our objective of profit maximization, having a strictly negative and strictly positive offset for the buyer's and seller's payments, respectively, would only deteriorate our objective. Therefore, we can restrict our focus to our definition of payments without any loss of generality.

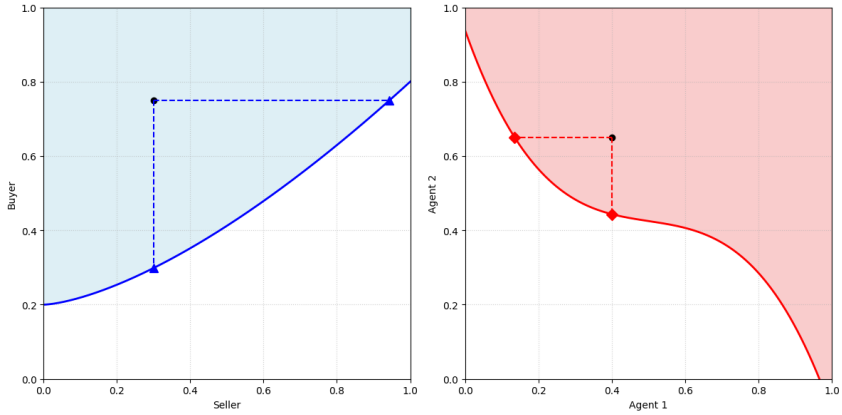

Figure 5: One bilateral trade mechanism (left) and one joint-ads mechanism (right), each with an allocated point and associated payments.

C.2 The Joint-Ads Problem

In the joint-ads problem, two buyers participate in an auction for a non-excludable[¶] specific good, and the goal of the mechanism designer is to maximize the revenue extracted from the trade.

This pair of agents is characterized by private valuations (v_1, v_2) in the $[0,1]^2$ square. The learner proposes a mechanism M to the agents, who then declare bids (b_1, b_2), with a mechanism being composed of an allocation region $A \subseteq [0,1]^2$ and two payment functions $p_1, p_2 : [0,1]^2 \to [0,1]$ that map bids to payments. The trade happens if $(b_1, b_2) \in A$ and the revenue of mechanism M when the bids are b_1 and b_2 is defined as

$$\mathsf{rev}(M, b_1, b_2) = p_1(b_1, b_2) + p_2(b_1, b_2).$$

Clearly, the allocation function and the payment function cannot be considered arbitrary. As in the bilateral trade problem, the agents behave strategically, meaning that they strive to maximize their own quasi-linear utility $u_i(b_1, b_2) = v_i \cdot \mathbb{1}_{\{(b_1,b_2)\in A\}} - p_i(b_1, b_2)$ for $i \in \{1, 2\}$. Note that, differently from the bilateral trade problem, here the two agents are symmetric, as they are both interested in purchasing a good. We require any proposed mechanism M to be DSIC and IR.

The family of mechanisms $\mathcal{M}^{JA}$ that respect these two properties can be easily characterized with the same techniques used for bilateral trade (and for any single-parameter mechanism design problem). This tells us that (i) an allocation rule is implementable if and only if it is monotone and (ii) for each such allocation rule there exists a unique Myerson payment rule that makes it a DSIC and IR mechanism. For this reason, we only talk about valuations in the rest of the appendix, as the agents are incentivized to behave truthfully.

Given that the agents are symmetric, the meaning of monotonicity and the definitions of payments are different from the ones in bilateral trade; in joint-ads, an allocation region is monotone if it is closed in the "north-east" direction, while payments are given by the "west" and "south" projection to the boundary of the allocation region. See Figure 5 for a visualization.

Definition 6 (Monotonicity and Payments for Joint-Ads)**.** *An allocation region $A \subseteq [0,1]^2$ is monotone for joint-ads if, for any $x = (x_1, x_2) \in A$ and $y = (y_1, y_2) \in [0,1]^2$ with $x_1 \leq y_1$ and $x_2 \leq y_2$, it holds that $y \in A$. Moreover, for any such monotone region A, Myerson payments are defined as follows:*

$$p_1(v_1, v_2) = \mathbb{1}_{\{(v_1,v_2)\in A\}} \cdot \min\{x \in [0,1] \mid (x, v_2) \in A\} \qquad \forall\, (v_1, v_2) \in [0,1]^2$$
$$p_2(v_1, v_2) = \mathbb{1}_{\{(v_1,v_2)\in A\}} \cdot \min\{y \in [0,1] \mid (v_1, y) \in A\} \qquad \forall\, (v_1, v_2) \in [0,1]^2$$

[§] $p(v_s, v_b) < 0$ and $q(v_s, v_b) > 0$ would violate the IR constraint.
[¶] Either both agents get the good, or none of the two gets it.

HEDGE

Input: finite action set A, learning rate $\eta \in (0, 1]$, time horizon T
Initialize $w_1(a) = 1$ for every $a \in A$
for $t = 1, \dots, T$ **do**
    Set $W_t = \sum_{a \in A} w_t(a)$ and $\pi_t(a) = {}^{w_t(a)}/_{W_t}$ for every $a \in A$
    Sample and play $a_t \sim \pi_t$
    Observe $r_t(a) \in [-1, 1]$ for every $a \in A$
    Set $w_{t+1}(a) = w_t(a) \exp(\eta \cdot r_t(a))$ for every $a \in A$

D The Hedge Algorithm

An important building block of our analysis is the HEDGE algorithm. For completeness, in this appendix we provide a brief description of the algorithm and the regret bound it achieves.

Consider the standard prediction with expert advice setting, on action set A, with possibly randomized reward functions $r_t : A \to [-1, 1]$ generated upfront by an oblivious adversary. As detailed in the pseudocode above, the HEDGE algorithm with learning rate η maintains a weight $w_t(a)$ for each action, defined as $w_t(a) = \exp\left(\eta \cdot \sum_{s=1}^{t-1} r_s(a)\right)$, and plays actions a_t by sampling proportionally to such weight. More precisely, the probability of choosing action a at time t is equal to $\pi_t(a) = {}^{w_t(a)}/_{W_t}$, where $W_t = \sum_{a'} w_t(a')$. In the main body we use the following standard bound, that we reprove here for completeness (for instance, see Theorem 2.3 in Arora et al. [2012]).

Lemma 5. *Consider the standard prediction with expert advice setting. The* HEDGE *algorithm with learning rate $\eta \in (0, 1]$ enjoys the following regret bound:*

$$\max_{a \in A} \sum_{t=1}^{T} \mathbb{E}\left[r_t(a) - r_t(a_t)\right] \leq \frac{\log |A|}{\eta} + \eta \cdot \mathbb{E}\left[\sum_{t=1}^{T} r_t(a_t)^2\right],$$

where the expectations are only with respect to the sampling of HEDGE*, with the inequality holding for every realization of the possibly randomized reward functions.*

Proof. The proof follows the standard pipeline, we just need to be a bit more careful about the scale of the rewards. The identity of the actual best action $a^\star$ only depends on the realized reward functions, and not on the actions chosen by the algorithm. Denote with W_t the sum of the weights of all the actions at the beginning of time t. We have:

$$\begin{aligned} \log(W_{T+1}) &\geq \log(w_{T+1}(a^\star)) && \text{(The weights are non-negative)} \\ &= \eta \sum_{t=1}^{T} r_t(a^\star). && (4) \end{aligned}$$

Consider now the generic time step t, we have:

$$\begin{aligned} \log \frac{W_{t+1}}{W_t} &= \log \sum_{a \in A} \frac{e^{\eta \sum_{s=1}^{t} r_s(a)}}{W_t} && \text{(By definition of } W_{t+1}) \\ &= \log \sum_{a \in A} \pi_t(a) e^{\eta \cdot r_t(a)} && \text{(By design of the algorithm)} \\ &\leq \log \sum_{a \in A} \pi_t(a) \left(1 + \eta \cdot r_t(a) + \eta^2 r_t^2(a)\right) && (e^x \leq 1 + x + x^2, \forall x < 1.79) \\ &= \log \left(1 + \eta \sum_{a \in A} \pi_t(a) r_t(a) + \eta^2 \sum_{a \in A} \pi_t(a) r_t^2(a)\right) \\ &\leq \eta \sum_{a \in A} \pi_t(a) r_t(a) + \eta^2 \sum_{a \in A} \pi_t(a) r_t^2(a), && (5) \end{aligned}$$

where the last inequality follows because $\log(1 + x) \leq x$ for all $x > -1$.

$$\log \frac{W_{T+1}}{W_1} = \log \prod_{t=1}^{T} \frac{W_{t+1}}{W_t} \qquad \text{(Telescopic argument)}$$

$$= \sum_{t=1}^{T} \log \frac{W_{t+1}}{W_t}$$
$$\leq \eta \sum_{t=1}^{T} \left[\sum_{a \in A} \pi_t(a) r_t(a) + \eta \sum_{a \in A} \pi_t(a) r_t^2(a) \right] \qquad \text{(By Equation (5))}$$

Noticing that $\log W_1 = \log|A|$, we get the claim. □

E Omitted Proof from Section 3

Lemma 1. *The cardinality of ε-grid mechanisms respects $2^{1/\varepsilon} \leq |\mathcal{M}_\varepsilon| \leq (2e)^{1/\varepsilon}$.*

Proof. Consider the directed graph G_ε whose nodes are given by the uniform grid V_ε, and whose edges are the sides of the tiles of $\mathcal{G}_\varepsilon$, pointing either from left to right (for horizontal edges) or bottom-up (for vertical ones). Each ε-grid mechanism is identified by a path on this graph going from $(0,0)$ to $(1,1)$, which — together with the left and upper sides of $[0,1]^2$ — encloses the allocation region. We then only need to count the number of paths of this type. Each path contains exactly $2/\varepsilon$ edges, equally divided into horizontal and vertical edges. All in all, we have $|\mathcal{M}_\varepsilon| = \binom{2/\varepsilon}{1/\varepsilon}$, so that

$$2^{1/\varepsilon} \leq |\mathcal{M}_\varepsilon| \leq (2e)^{1/\varepsilon},$$

where we use the inequalities $(n/k)^k \leq \binom{n}{k} \leq (en/k)^k$. □

F From Joint-Ads to Bilateral Trade: Reduction and Corollaries

In this Appendix, we formally reduce the joint-ads problem to bilateral trade. In particular, in Theorem 2 we argue that any learning algorithm for (profit maximization in) bilateral trade achieving some regret bound $\rho(\sigma, T)$ can be used to derive a learning algorithm for joint-ads suffering $O(\rho(\sigma, T))$ regret.

The main ingredient for the reduction is simple: an affine transformation $f : [0,1]^2 \to [0, 1/2] \times [1/2, 1]$ defined as follows:

$$f(x,y) = \left(\frac{1-x}{2}, \frac{1+y}{2} \right).$$

It is immediate to see that f is bijective from $[0,1]^2$ to $[0, 1/2] \times [1/2, 1]$.

Theorem 2 (A Reduction from Joint-Ads to Bilateral Trade)**.** *Let $\mathcal{A}_{\text{BT}}$ be a learning algorithm for profit-maximization in bilateral trade against the σ-smooth adversary which enjoys a regret bound $\rho(\sigma, T)$, with σ possibly unknown to the algorithm. Then we can construct an algorithm $\mathcal{A}_{\text{JA}}$ for joint-ads against the σ'-smooth adversary which enjoys a regret bound $O(\rho(\sigma, T))$, where $\sigma = \sigma'/4$.*

Proof. Given a learning algorithm $\mathcal{A}_{\text{BT}}$ for bilateral trade, we show how to derive $\mathcal{A}_{\text{JA}}$ for joint ads which enjoys the same regret rate. We construct $\mathcal{A}_{\text{JA}}$ by running $\mathcal{A}_{\text{BT}}$ on a suitable instance of bilateral trade. The reduction proceeds in two stages at each time step.

At each time t, we query mechanism M_{BT}^t from $\mathcal{A}_{\text{BT}}$, and construct a mechanism M_{JA}^t by *clipping* it to the square of interest $Q = [0, 1/2] \times [1/2, 1]$ (see first plot in Figure 4 in the main body), and applying the inverse of f defined above. Formally, if A_{BT}^t is the allocation region of M_{BT}^t, we define the allocation region A_{JA}^t of M_{JA}^t as follows:

$$A_{\text{JA}}^t = \left\{ (x', y') | f(x', y') \in A_{\text{BT}}^t \cap Q \right\}. \tag{6}$$

Then, we observe the valuation (v_1^t, v_2^t) generated from the joint-ads adversary, and feed $(v_s^t, v_b^t) = f(v_1^t, v_2^t)$ back to the bilateral trade algorithm $\mathcal{A}_{\text{BT}}$. The rest of the proof is devoted to proving that this is indeed a reduction with the desired properties.

Fix any σ'-smooth instance generated by the joint ads adversary. Since the bilateral trade algorithm sees the valuations generated by the joint-ads-σ'-smooth adversary mapped back through f, then a

simple calculation yields that $\mathcal{A}_{\text{BT}}$ is playing against a σ-smooth distribution, and thus is guaranteed the following regret bound by hypothesis:

$$\sup_{M_{\text{BT}} \in \mathcal{M}_{\text{BT}}} \sum_{t=1}^{T} \mathbb{E}\left[\mathsf{profit}_t(M_{\text{BT}}) - \mathsf{profit}_t(M_{\text{BT}}^t)\right] \leq \rho(\sigma, T). \tag{7}$$

Note, the $\mathsf{profit}_t(\cdot)$ functions compute the profit collected by a bilateral trade mechanism under valuations (v_s^t, v_b^t) constructed as in the reduction.

The next step is crucial: we need to relate the revenue $\mathsf{rev}_t(M_{\text{JA}}^t)$ extracted by the mechanism M_{JA}^t in the joint ads instance with the profit $\mathsf{profit}_t(M_{\text{BT}}^t)$ extracted by the mechanism M_{BT}^t in the corresponding bilateral trade instance.

Claim 2. *The following inequality is verified:* $\mathsf{profit}_t(M_{\text{BT}}^t) \leq \frac{1}{2}\mathsf{rev}_t(M_{\text{JA}}^t)$.

Proof of Claim 2. First, we observe that by construction (v_s^t, v_b^t) belongs to Q. Therefore (v_1^t, v_2^t) belongs to the allocation region of M_{JA}^t if and only if the corresponding bilateral-trade-valuation (v_s^t, v_b^t) belongs to the allocation region of M_{BT}^t. We are then sure that when the trade does not happen in the bilateral trade instance, the same happens also in the joint ads one, and vice-versa.

Consider now the case in which both $(v_1^t, v_2^t) \in A_{\text{JA}}^t$ and the corresponding (v_s^t, v_b^t) belongs to A_{BT}^t (and also to Q by construction). We have the following:

$$\begin{aligned}
\mathsf{profit}_t(M_{\text{BT}}^t) &= \min\{y \mid (v_s^t, y) \in A_{\text{BT}}^t\} - \max\{x \mid (x, v_b^t) \in A_{\text{BT}}^t\} \\
&\leq \min\{y \mid (v_s^t, y) \in A_{\text{BT}}^t \cap Q\} - \max\{x \mid (x, v_b^t) \in A_{\text{BT}}^t \cap Q\} \\
&= \min\{y \mid f^{-1}(v_s^t, y) \in A_{\text{JA}}^t\} - \max\{x \mid f^{-1}(x, v_b^t) \in A_{\text{JA}}^t\} \quad \text{(As } f \text{ is bijective)} \\
&= \min\{y \mid (v_1^t, 2y-1) \in A_{\text{JA}}^t\} - \max\{x \mid (1-2x, v_2^t) \in A_{\text{JA}}^t\} \\
&= \min\left\{\frac{1+y'}{2} \mid (v_1^t, y') \in A_{\text{JA}}^t\right\} - \max\left\{\frac{1-x'}{2} \mid (x', v_2^t) \in A_{\text{JA}}^t\right\} \\
&= \frac{1}{2}\min\left\{y' \mid (v_1^t, y') \in A_{\text{JA}}^t\right\} + \frac{1}{2}\min\left\{x' \mid (x', v_2^t) \in A_{\text{JA}}^t\right\} \\
&= \frac{1}{2}\mathsf{rev}_t(M_{\text{JA}}^t).
\end{aligned} \tag{8}$$

This concludes the proof. □

Consider the best mechanism $M_{\text{JA}}^\star$ for the σ'-smooth instance of joint ads[‖], and denote with $M_{\text{BT}}^\star$ the mechanism for bilateral trade constructed analogously to Equation (6). We have the following result.

Claim 3. *The following equality holds for any t:* $\mathsf{profit}_t(M_{\text{BT}}^\star) = \frac{1}{2}\mathsf{rev}_t(M_{\text{JA}}^\star)$.

Proof of Claim 3. The claim follows by exactly the same arguments of Claim 2. The only difference is that, by construction, the allocation region of $M_{\text{BT}}^\star$ is contained in Q, therefore the only inequality of Equation (8) holds with equality. □

We have all the ingredients to bound the regret of $\mathcal{A}_{\text{JA}}$:

$$\begin{aligned}
&\sum_{t=1}^{T} \mathbb{E}\left[\mathsf{rev}_t(M_{\text{JA}}^\star) - \mathsf{rev}_t(M_{\text{JA}}^t)\right] \\
&\quad\leq 2\sum_{t=1}^{T} \mathbb{E}\left[\mathsf{profit}_t(M_{\text{BT}}^\star) - \mathsf{profit}_t(M_{\text{BT}}^t)\right] && \text{(By Claims 2 and 3)} \\
&\quad\leq 2\rho(\sigma, T), && \text{(By Equation (7))}
\end{aligned}$$

which concludes the proof of the theorem. □

[‖]Also for joint ads one can prove that the sup in the regret definition is attained, but this proof carries over with minimal modifications by considering some small precision ε and the corresponding ε-optimal mechanism.

The reduction has two immediate corollaries. First, we can apply directly the reduction to translate the regret rate of HIER-MECH proved in Theorem 1 to the joint ads problem.

Corollary 2 (An upper bound for joint ads)**.** *Consider the problem of revenue maximization in repeated joint ads against a σ-smooth adversary, with σ possibly unknown. There exists a learning algorithm which suffers regret at most $\frac{160}{\sigma}\sqrt{T}\log T$.*

As a second corollary, we can apply the reduction in the other direction, and claim that no algorithm for bilateral trade can hope to achieve better than $\Omega(\sqrt{T})$ regret *given* the lower bound of the same rate for joint ads [see Theorem 6.2 of Aggarwal et al., 2024].

Corollary 1 (A lower bound for bilateral trade)**.** *Consider the problem of profit maximization in repeated bilateral trade against a σ-smooth adversary, for $\sigma \in (0, 1/48)$. Then, for any learning algorithm $\mathcal{A}$, the following lower bound holds: $R_T(\mathcal{A}) \geq \frac{3}{128}\sqrt{T}$.*