Near-optimal Swap Regret Minimization for Convex Losses

Lunjia Hu LUNJIA@ALUMNI.STANFORD.EDU
Northeastern University

Jon Schneider JSCHNEI@GOOGLE.COM
Google

Yifan Wu YIFAN.WU2357@GMAIL.COM
Microsoft Research

Editors: Steve Hanneke and Tor Lattimore

Abstract

We give a randomized online algorithm that guarantees near-optimal $\widetilde{O}(\sqrt{T})$ expected swap regret against any sequence of T adaptively chosen Lipschitz convex losses on the unit interval. This improves the previous best bound of $\widetilde{O}(T^{2/3})$ and answers an open question of Fishelson et al. (2025b). In addition, our algorithm is efficient: it runs in $\mathsf{poly}(T)$ time. A key technical idea we develop to obtain this result is to discretize the unit interval into bins at multiple scales of granularity and simultaneously use all scales to make randomized predictions, which we call multi-scale binning and may be of independent interest. A direct corollary of our result is an efficient online algorithm for minimizing the calibration error for general elicitable properties. This result does not require the Lipschitzness assumption of the identification function needed in prior work, making it applicable to median calibration, for which we achieve the first $\widetilde{O}(\sqrt{T})$ calibration error guarantee.

Keywords: Swap regret; calibration; property elicitation; sequential forecasting

1. Introduction

Swap regret minimization is a fundamental problem in the field of online learning, with a recent uptick in interest due to its many applications to calibration (Foster and Vohra, 1997, 1998), fairness (Hébert-Johnson et al., 2018; Gopalan et al., 2023), and learning in games (Deng et al., 2019b). Like the standard notion of external regret, swap regret measures the suboptimality of a sequence of actions taken by an online learner against an adversary; however, whereas minimizing external regret only guarantees that there is no single action that performs better in hindsight, minimizing swap regret provides the much stronger guarantee that there is no consistent transformation of the learner's actions (e.g., substituting action A every time the learner played action B) that outperforms the learner.

To this date, the majority of results on swap regret minimization focus on settings where the losses incurred by the learner are *linear* functions of the learner's actions (for example, in the setting originally studied by Blum and Mansour (2007), every round the learner picks a mixed strategy $p \in \Delta_n$ over n actions, the adversary picks an n-dimensional loss $\ell \in [-1, 1]^n$, and the learner incurs the linear loss $\langle p, \ell \rangle$). Recently, Fishelson et al. (2025b) initiated the study of swap regret minimization against sequences of *non-linear* convex losses, showing that many important problems (L_2 calibration, minimization of swap regret in structured games) can be conveniently framed in this language.

Formally, Fishelson et al. (2025b) introduce the following problem. A Learner interacts with an Adversary for T rounds. In each round $t \in [T]$, the Learner begins by choosing a distribution ρ_t supported on $[0,1]$. The Adversary then reveals a 1-Lipschitz convex loss function $\ell_t : [0,1] \to \mathbb{R}$. The Learner then samples an action p_t from ρ_t and incurs loss $\ell_t(p_t)$. The goal of the Learner is to minimize their expected swap regret $\mathbb{E}[\mathsf{SR}(p_{1:T}; \ell_{1:T})]$, where

$$\mathsf{SR}(p_{1:T}; \ell_{1:T}) := \sup_{\sigma:[0,1]\to[0,1]} \left(\sum_{t=1}^{T} \ell_t(p_t) - \sum_{t=1}^{T} \ell_t(\sigma(p_t)) \right). \tag{1}$$

If all the losses ℓ_t are linear functions in p, then one can attain $\widetilde{O}(\sqrt{T})$ swap regret by simply running a standard swap-regret minimization algorithm over the two endpoints of the Learner's action set $[0,1]$ (in particular, any intermediate action $p \in (0,1)$ can be replaced by a distribution playing 1 with probability p and 0 with probability $1-p$ without changing the Learner's expected loss). Similarly, if the losses ℓ_t are strongly convex, Fishelson et al. (2025b) show that it is possible to obtain $\widetilde{O}(T^{1/3})$ (pseudo[1]-)swap regret bounds, by taking advantage of $O(\log T)$ external regret bounds for such losses.

Interestingly, despite the $\widetilde{O}(\sqrt{T})$ swap regret bound for linear losses and the $\widetilde{O}(T^{1/3})$ swap regret bound for smooth, strongly convex losses, the best known attainable for arbitrary (bounded and Lipschitz) convex losses is $\widetilde{O}(T^{2/3})$. This is worse than either of the two above bounds, and follows from simply discretizing the Learner's action set and running a classical swap-regret algorithm over the discretization. This is particularly counter-intuitive because minimizing external regret for convex losses is *no harder* than minimizing external regret for linear losses – it suffices to simply minimize external regret with respect to the linearizations of loss function formed by the subgradients of ℓ_t at the points p_t played by the Learner. Unfortunately, this reduction does not extend to swap regret minimization[2]. Fishelson et al. (2025b) pose as an open question whether it is possible to improve this bound to match the $\widetilde{O}(\sqrt{T})$ bound for linear losses.

1.1. Our Contributions

We positively resolve this open question, providing an efficient algorithm achieving $\widetilde{O}(\sqrt{T})$ swap regret against an arbitrary adaptive sequence of (bounded and Lipschitz) convex losses.

Problem 1 (Swap regret minimization for convex losses) *In each round $t = 1, \ldots, T$,*

1. *predictor (i.e. learner) chooses a distribution ρ_t on $[0,1]$;*
2. *adversary reveals a 1-Lipschitz convex loss function $\ell_t : [0,1] \to \mathbb{R}$;*

1. Fishelson et al. (2025b) study a slightly weaker definition of swap-regret than defined above, where the expectation over sampling from ρ_t takes place within the supremum of (1) (i.e., for any specific swap function σ they guarantee $\widetilde{O}(T^{1/3})$ expected swap-regret). This was later strengthened by Luo et al. (2025b), who proved an $\widetilde{O}(T^{1/3})$ expected swap-regret bound for sequences of smooth convex losses arising from online calibration instances.
2. The subtlety here is that the linearization of ℓ_t depends on the specific action p_t sampled from ρ_t. The reduction for external regret minimization sidesteps this because the Learner can always play a deterministic strategy (i.e., a singleton ρ_t) – in contrast, deterministic algorithms cannot guarantee sublinear swap regret. Indeed, against any deterministic predictor, after seeing p_t the adversary may play $\ell_t(p) = p$ if $p_t > 1/2$ and $\ell_t(p) = 1-p$ otherwise. The predictor incurs loss at least $1/2$ every round, while the swap function that sends predictions above $1/2$ to 0 and predictions at most $1/2$ to 1 has zero loss, giving swap regret at least $T/2$. This argument uses that the realized deterministic prediction is known before ℓ_t is chosen; it does not apply to randomized predictors.

3. *prediction* $p_t \in [0,1]$ *is drawn from* ρ_t*, and predictor incurs loss* $\ell_t(p_t)$*.*

The predictor's goal is to minimize the expected swap regret $\mathbb{E}[\mathsf{SR}(p_{1:T};\ell_{1:T})]$ *defined in* (1).

Theorem 2 *For every positive integer* $T \geq 2$*, there exists a prediction strategy for Problem 1 that guarantees*

$$\mathbb{E}[\mathsf{SR}(p_{1:T};\ell_{1:T})] = O(\sqrt{T}\log T).$$

Moreover, there exists such a prediction strategy that runs in $\mathsf{poly}(T)$ *time.*

Theorem 2 is optimal up to the $O(\log T)$ factor because there is a simple way to choose the losses $\ell_{1:T}$ to incur $\Omega(\sqrt{T})$ expected swap regret against any prediction strategy (Lemma 4). In addition to the expectation bound in Theorem 2, our prediction algorithm also guarantees a high-probability bound: given any $\delta \in (0, 1/T]$, the algorithm achieves $O\left(\sqrt{(T\log T)\log(1/\delta)}\right)$ swap regret with probability at least $1-\delta$ (Lemma 14).

Feedback model. Our main result is for the full-information setting, where the revealed loss can be queried after the prediction distribution is chosen. A bandit version, where the learner only observes $\ell_t(p_t)$ or another partial-information signal, is a natural direction for future work. If the losses are additionally bounded in $[-1,1]$, discretizing $[0,1]$ and applying a standard finite-action bandit algorithm for swap or internal regret already gives sublinear swap regret, but the optimal rate in this feedback model remains unclear.

1.1.1. IMPLICATIONS TO CALIBRATION FOR ELICITABLE PROPERTIES

As a motivating application of our results, we describe how Theorem 2 leads to improved algorithms for the problem of *calibrated forecasting of elicitable properties*. In particular, we show that such calibration measures of common properties (such as mean, median, and q-quantile) can often be written as swap regrets of scoring functions that are Lipschitz and convex, allowing Theorem 2 to recover the first $\widetilde{O}(T^{1/2})$ calibration bounds for such properties.

Proper Scoring Rules and Elicitable Properties. Many prediction tasks are to estimate a *statistical property* of an unknown outcome distribution. Formally, let τ be a distribution over outcomes Y, and let Γ be a (possibly set-valued) property mapping that assigns to each τ a subset $\Gamma(\tau) \subseteq [0,1]$ of optimal reports (e.g., probabilities, quantiles, or other statistics). A predictor outputs a report $p \in [0,1]$, after which an outcome $y \sim \tau$ is realized.

A *scoring function* (or loss) is a function $S : [0,1] \times Y \to \mathbb{R}$ that evaluates the report p when the realized outcome is y. The score S is *proper*, if it *elicits* the property Γ: for every distribution τ over Y, the set of Bayes-optimal reports under τ coincides with $\Gamma(\tau)$:

$$\arg\min_{p\in[0,1]} \mathbb{E}_{y\sim\tau}[S(p,y)] \;=\; \Gamma(\tau). \tag{2}$$

A property Γ is called *elicitable* if there exists some scoring function S that elicits it. Assuming the outcome space is $Y = [0,1]$, the following scoring functions elicit the mean, the median, and the q-quantile:

$$S_{\mathsf{mean}}(p,y) = (p-y)^2, \tag{3}$$

$$S_{\mathsf{median}}(p,y) = |p-y|, \tag{4}$$

$$S_{q\text{-quantile}}(p, y) = \mathbb{I}\left[y \geq p\right] q \cdot (p - y) + \mathbb{I}\left[y \leq p\right] (1 - q) \cdot (y - p) \tag{5}$$

Note that all these scoring functions $S(\cdot, y)$ are convex and $O(1)$-Lipschitz for any fixed y.

Calibration Error for Elicitable Properties. Given a scoring function $S : [0, 1] \times Y \to \mathbb{R}$, we define the sub-optimality of $p \in [0, 1]$ w.r.t. distribution τ on Y as follows:

$$\mathsf{subopt}_\tau(p) := \mathbb{E}_{y\sim\tau} S(p, y) - \inf_{p^* \in [0,1]} \mathbb{E}_{y\sim\tau} S(p^*, y) \geq 0.$$

When Γ is the property elicited by S as in (2), we have $\mathsf{subopt}_\tau(p) = 0$ if and only if $p \in \Gamma(\tau)$. For sequences $p_1, \ldots, p_T \in [0, 1]$ and $y_1, \ldots, y_T \in Y$, we define the calibration error CAL_S as follows:

$$\mathsf{CAL}_S(p_{1:T}; y_{1:T}) := \sum_{p\in P} n_p \cdot \mathsf{subopt}_{\tau_p}(p), \tag{6}$$

where $P = \{p_1, \ldots, p_T\}$ is the set of distinct predictions, $n_p = |\{t \in [T] : p_t = p\}|$ is the number of rounds t such that $p_t = p$, and τ_p is the distribution of y_t when t is sampled uniformly from $\{t \in [T] : p_t = p\}$.

We use $\mathsf{CAL}_{\mathsf{median}}$, $\mathsf{CAL}_{\mathsf{mean}}$, and $\mathsf{CAL}_{q\text{-quantile}}$ to denote the swap regret defined by S_{median}, S_{mean}, and $S_{q\text{-quantile}}$, respectively. In particular, $\mathsf{CAL}_{\mathsf{mean}}$ is the standard ℓ_2 calibration error:

$$\mathsf{CAL}_{\mathsf{mean}}(p_1, \ldots, p_T; y_1, \ldots, y_T) \ = \ \sum_{p\in P} n_p \Big(p - \mathbb{E}_{y\sim\tau_p}[y]\Big)^2. \tag{7}$$

The calibration error CAL_S in (6) can be naturally formulated as swap regret:

$$\mathsf{CAL}_S(p_{1:T}; y_{1:T}) = \sup_{\sigma:[0,1]\to[0,1]} \left(\sum_{t=1}^{T} S(p_t, y_t) - \sum_{t=1}^{T} S(\sigma(p_t), y_t) \right).$$

Thus our Theorem 2 immediately implies the following corollary:

Corollary 3 *Let $S : [0, 1] \times Y \to \mathbb{R}$ be a scoring function such that $S(p, y)$ is a convex and 1-Lipschitz function of $p \in [0, 1]$ for every $y \in Y$. For every positive integer T, there exists a randomized prediction strategy that guarantees*

$$\mathbb{E}[\mathsf{CAL}_S(p_{1:T}; y_{1:T})] = O(\sqrt{T} \log T).$$

As a result, there exists a randomized prediction strategy that guarantees $O(\sqrt{T} \log T)$ for $\mathsf{CAL}_{\mathsf{median}}$ and $\mathsf{CAL}_{q\text{-quantile}}$[3].

The guarantee above is near-optimal for median calibration.

Lemma 4 *Suppose $y_1, \ldots, y_T$ are drawn i.i.d. from the Bernoulli distribution with mean $1/2$. For any prediction algorithm, $\mathbb{E}[\mathsf{CAL}_{\mathsf{median}}(p_{1:T}, y_{1:T})] = \Omega(\sqrt{T})$.*

3. $\mathsf{CAL}_{\mathsf{mean}}$ corresponds to a *strongly* convex and smooth loss function, where $\widetilde{O}\left(T^{1/3}\right)$ can be guaranteed (Fishelson et al., 2025b).

Lemma 4 can be proved by noting that $\mathbb{E}[|(y_1 + \cdots + y_T) - T/2|] = \Omega(\sqrt{T})$, implying that $\mathbb{E}[\inf_{p\in[0,1]} \sum_{t=1}^{T} S(p, y_t)] \le T/2 - \Omega(\sqrt{T})$ for scoring rule $S = S_{\mathsf{median}}$. However, we always have $\mathbb{E}[S(p_t, y_t)] = 1/2$. Therefore,

$$\mathbb{E}[\mathsf{CAL}_{\mathsf{median}}(p_{1:T}, y_{1:T})] \ge \mathbb{E}\left[\sum_{t=1}^{T} S(p_t, y_t) - \inf_{p\in[0,1]} \sum_{t=1}^{T} S(p, y_t)\right] \ge \Omega(\sqrt{T}).$$

We note that results in prior work on calibration for elicitable properties (e.g. Noarov and Roth (2023); Hu et al. (2025)) do not apply to median calibration unless additional Lipschitzness assumptions about the distribution of y are satisfied. Previous work also considers a different definition of calibration error for elicitable properties. See Appendix B.3 for a detailed discussion.

1.2. Technical Overview

We give a high-level explanation for how we prove Theorem 2, highlighting the core ideas. Our main technical innovation lies in the multi-scale binning idea we discuss below.

V-shaped Decomposition. In each round t, the loss function $\ell_t : [0, 1] \to \mathbb{R}$ we receive from the adversary is convex and 1-Lipschitz. It is a commonly used fact that such losses can be written as a convex combination of *V-shaped losses* up to an additive constant: there exists a distribution φ_t on $[0, 1]$ and a constant $C_t \in \mathbb{R}$ such that $\ell_t(p) = \mathbb{E}_{v\sim\varphi_t}|p-v|+C_t$ for every $p \in [0, 1]$. This V-shaped decomposition has been established and used in many prior works (Li et al., 2022; Kleinberg et al., 2023; Gopalan et al., 2024; Hu and Wu, 2024). We formally state this result in Lemma 5 and include a proof for completeness.

This decomposition suggests that we focus on the *base* V-shaped losses of the form $\ell_t(p) = |p - v_t|$ parameterized by $v_t \in [0, 1]$. In this subsection, we will assume that every loss ℓ_t has this form. This allows us to reason about points $v_t \in [0, 1]$ rather than functions $\ell_t : [0, 1] \to \mathbb{R}$.

Non-constructive Proof via the Minimax Theorem. Problem 1 can be viewed as a zero-sum game between the predictor and the adversary. The predictor plays a mixed strategy that determines the prediction $p_t \in [0, 1]$ in every round based on the transcript of previous rounds. After seeing the predictor's mixed strategy, the adversary plays a strategy that determines the loss ℓ_t in each round also based on the transcript of previous rounds. The value of the game is the expected swap regret when both players play optimally, with the predictor minimizing the expected swap regret and the adversary maximizing it. If we ignore the running time guarantee in Theorem 2, the theorem is equivalent to the claim that the value of this zero-sum game is $O(\sqrt{T} \log T)$.

The minimax theorem tells us that the value of a two-player zero-sum game is invariant to who plays first. We thus consider an order-reversed game, where the adversary first plays a mixed strategy to choose ℓ_t, and after seeing this mixed strategy, the predictor then decides their strategy of choosing p_t. It is easier to construct a good predictor for this order-reversed game because the predictor knows more information than in the original game. In Section 3 we design a prediction strategy that guarantees $O(\sqrt{T} \log T)$ expected swap regret for the order-reversed game, implying a non-constructive version of Theorem 2 that does not have the running time guarantee. To fully prove Theorem 2, we apply a multi-objective learning framework to reset the order of play in an efficient way (Appendix A). We further explain how we prove both the non-constructive and original versions of Theorem 2 below.

Binning. When minimizing swap regret over a continuous prediction space $[0,1]$, it is natural to restrict the predictions to a finite subset $B \subseteq [0,1]$ of bins. This is beneficial because when two predictions p_t and $p_{t'}$ are very close, making them exactly equal ($p_t = p_{t'}$) adds a constraint $\sigma(p_t) = \sigma(p_{t'})$ to the swap function σ in (1), which is helpful for reducing the swap regret. Binning is also helpful when we apply the minimax theorem which requires the action spaces of both players to be finite.

Once we restrict all predictions p_t to a finite set $B \subseteq [0,1]$, we can decompose the swap regret in (1) by contribution from each bin $b \in B$. Define $T_b := \{t \in [T] \mid p_t = b\}$ for every $b \in B$, then

$$\mathsf{SR}(p_{1:T}; \ell_{1:T}) = \sum_{b \in B} \left(\sum_{t \in T_b} \ell_t(b) - \sum_{t \in T_b} \ell_t(s_b) \right), \quad \text{where } s_b \in \operatorname*{arg\,min}_{s \in [0,1]} \sum_{t \in T_b} \ell_t(s). \tag{8}$$

Truthful Predictor. As we will discuss soon, it is a non-trivial task to choose the set of bins B appropriately. Once B is chosen, a natural prediction strategy for the order-reversed game is to make *truthful* predictions.

Recall that the adversary plays first in the order-reversed game and they can use a mixed strategy. In each round $t = 1, \ldots, T$, the adversary first provides a distribution π_t of loss functions ℓ. Based on π_t, the predictor chooses $p_t \in B$. After that, a loss function ℓ_t is drawn from π_t and the predictor incurs loss $\ell_t(p_t)$.

Knowing the loss distribution π_t, a truthful predictor simply chooses $p_t \in B$ to minimize the expected loss $\bar{\ell}_t(p_t)$, where $\bar{\ell}_t := \mathbb{E}_{\ell \sim \pi_t}[\ell]$. We analyze the swap regret of this truthful prediction strategy as follows.

Sampling error and rounding error. Let us focus on an arbitrary fixed bin $b \in B$. We define loss functions averaged over rounds $t \in T_b$:

$$\bar{\ell} := \frac{1}{|T_b|} \sum_{t \in T_b} \bar{\ell}_t, \quad \text{and} \quad \hat{\ell} := \frac{1}{|T_b|} \sum_{t \in T_b} \ell_t.$$

By (8), the swap regret contribution SR_b from bin b is the following:

$$\frac{1}{|T_b|} \mathsf{SR}_b = \frac{1}{|T_b|} \left(\sum_{t \in T_b} \ell_t(b) - \sum_{t \in T_b} \ell_t(s_b) \right) = \hat{\ell}(b) - \hat{\ell}(s_b), \quad \text{where } s_b \in \operatorname*{arg\,min}_{s \in [0,1]} \hat{\ell}(s).$$

Analogously to the definition of s_b, we define

$$\bar{s}_b \in \operatorname*{arg\,min}_{s \in [0,1]} \bar{\ell}(s).$$

The truthful prediction strategy ensures that b is the minimizer of $\bar{\ell}_t(b')$ over $b' \in B$ for every $t \in T_b$. Therefore,

$$b \in \operatorname*{arg\,min}_{b' \in B} \bar{\ell}(b').$$

Both b and $\bar{s}_b$ are minimizers of the function $\bar{\ell}$. The difference is that b comes from a restricted set B, whereas $\bar{s}_b$ is chosen from the full domain $[0,1]$. Thus we view the difference between b and $\bar{s}_b$ as *rounding error* caused by binning.

The swap regret SR_b contributed by bin b can be decomposed as follows:

$$\frac{1}{|T_b|}\mathsf{SR}_b = \hat{\ell}(b) - \hat{\ell}(s_b) = \Big(\hat{\ell}(b) - \hat{\ell}(s_b)\Big) - \Big(\bar{\ell}(b) - \bar{\ell}(s_b)\Big) + \Big(\bar{\ell}(b) - \bar{\ell}(s_b)\Big)$$
$$\le \underbrace{\Big(\hat{\ell}(b) - \hat{\ell}(s_b)\Big) - \Big(\bar{\ell}(b) - \bar{\ell}(s_b)\Big)}_{\text{Sampling error between } \hat{\ell} \text{ and } \bar{\ell}} + \underbrace{\Big(\bar{\ell}(b) - \bar{\ell}(\bar{s}_b)\Big)}_{\text{Rounding error between } b \text{ and } \bar{s}_b} .$$

There are two sources of error: 1) sampling error between $\hat{\ell}$ and $\bar{\ell}$, and 2) rounding error between b and $\bar{s}_b$. Roughly speaking, increasing the number of bins reduces the rounding error, but it results in fewer data points per bin, causing the sampling error to increase. Our choice of the bins B should minimize the sum of the two types of error.

Challenge with Fixed Binning. The key technical challenge is that no fixed choice of B can guarantee $\widetilde{O}(\sqrt{T})$ swap regret for the truthful prediction strategy.

Consider the case where all π_t's are the same singleton distribution on a single V-shaped loss function $\ell(p) = |p - v|$. This distribution is deterministic, so we always have $\hat{\ell} = \bar{\ell}$ and there is no sampling error. However, the **rounding error** can be very large: $\bar{\ell}(\bar{s}_b) = 0$ is the minimum value of $\bar{\ell}$, so the rounding error $\bar{\ell}(b) - \bar{\ell}(\bar{s}_b) = |b - v|$ is the absolute difference between b and v. To ensure $\widetilde{O}(\sqrt{T})$ swap regret, we need $|b - v| = \widetilde{O}(\sqrt{T})/T = 1/\widetilde{\Omega}(\sqrt{T})$, but in the worst case, v can fall at the midpoint of two adjacent bins, so the gap between any two adjacent bins should be at most $1/\widetilde{\Omega}(\sqrt{T})$. In particular, there needs to be at least $\widetilde{\Omega}(\sqrt{T})$ bins in total.

Now suppose there are m bins in B, and assume without loss of generality that at least $m/2$ of these bins b fall in the first half $[0, 1/2]$ of the unit interval. Consider the case where for every such bin $b \le 1/2$ and every $t \in T_b$, we choose π_t to be a near-uniform distribution on two losses $\ell(p) = |p - b|$ and $\ell'(p) = |p - (b + 1/2)|$ assigning slightly larger probability mass $1/2 + \varepsilon$ to ℓ than the probability mass $1/2 - \varepsilon$ on ℓ'. The expected loss $\bar{\ell} := (1/2 + \varepsilon)\ell + (1/2 - \varepsilon)\ell'$ is indeed minimized at b, so the truthful predictor predicts $p_t = b$ and there is no rounding error. However, the **sampling error** can be very large. The loss $\hat{\ell}$ can be decomposed as $\eta\ell + (1 - \eta)\ell'$ with η being the head frequency of $|T_b|$ independent tosses of a coin with bias $1/2 + \varepsilon$. As $\varepsilon \to 0$, with probability $\Theta(1)$, we have $\eta \le 1/2 - \Omega(1/\sqrt{|T_b|})$, in which case the minimizer s_b of $\hat{\ell}$ is $b + 1/2$ instead of b, and the difference $\hat{\ell}(b) - \hat{\ell}(s_b)$ is $\Omega(1/\sqrt{|T_b|})$. Therefore, the expected swap regret contribution from bin b is $\mathsf{SR}_b = \Omega(\sqrt{T_b})$. Assuming that the bins have roughly equal sizes $|T_b| \approx T/m$, the total swap regret is $\Omega(m \cdot \sqrt{T/m}) = \Omega(\sqrt{Tm})$. Thus we need as few as $m = \mathrm{polylog}(T)$ bins to ensure $\widetilde{O}(\sqrt{T})$ swap regret, contradicting with the earlier requirement of having at least $\widetilde{\Omega}(\sqrt{T})$ bins.

Multi-scale Binning. Our main technical contribution is to use *multi-scale binning* to address the challenge above. The idea is to adapt the bins B to the distribution π_t in each round t. For example, in the case above where π_t is deterministic on $\ell(p) = |p - v|$, we use as many as $\widetilde{\Omega}(\sqrt{T})$ fine-grained bins to reduce the rounding error. Instead, when π_t is the uniform distribution on $\ell(p) = |p - v|$ and $\ell'(p) = |p - (v + 1/2)|$, we use as few as $O(1)$ coarse-grained bins to reduce the sampling error.

In general, the granularity of binning needs to be carefully determined based on the distribution π_t. To that end, we introduce the notion of *width* w_t of a distribution π_t and choose the binning granularity to match w_t. Assuming that every loss $\ell \sim \pi_t$ is a V-shaped loss $\ell(p) = |p - v|$, we interpret π_t as a distribution of $v \in [0, 1]$. Intuitively, when π_t is deterministic and supported on a single $v \in [0,1]$, it has small width and we should use fine-grained bins, whereas when π_t is uniform on $\{v, v+1/2\}$, it has large width and we should use coarse-grained bins. Our exact definition of width for a general distribution on $[0,1]$ is based on a carefully constructed equation about the distance between two quantiles of π_t (Lemma 8). This definition allows us to optimally balance the rounding and sampling error and show the $O(\sqrt{T}\log T)$ expected swap regret bound for the order-reversed game. Since the number of predictions that fall in each bin can differ substantially, our full proof in Section 3 involves a case analysis that handles heavy and light bins separately.

Efficient Algorithm via Multi-Objective Learning. To convert our multi-scale binning predictor for the order-reversed game to an efficient predictor for the original Problem 1, we reduce Problem 1 to a multi-objective learning task based on the concentration inequalities we used when proving the swap regret guarantee for the order-reversed game. Each concentration inequality corresponds to a constraint or an "objective" in multi-objective learning, and our goal is to make predictions to satisfy all of them. This is inspired by a line of work that solves similar problems (e.g. online calibration, multicalibration, and omniprediction) using multi-objective learning, such as the Blackwell approachability framework of Abernethy et al. (2011) and the *Adversary-Moves-First (AMF)* framework of Lee et al. (2022). A key ingredient needed to solve multi-objective learning tasks is an *expert algorithm* that computes a convex combination of the "objectives" in every round. To obtain our desired swap regret bound, we use the MsMwC algorithm of Chen et al. (2021) which provides a stronger guarantee than the usual multiplicative weights algorithm, following Noarov et al. (2025); Hu and Wu (2024); Hu et al. (2025). Our final efficient predictor makes randomized predictions over multiple bins across different scales at every round. See Appendix A for our full solution to Problem 1.

1.3. Related Work

We defer additional related work to Appendix B.

2. Preliminaries

The following standard lemma decomposes any 1-Lipschitz convex loss $\ell : [0,1] \to \mathbb{R}$ into a convex combination of V-shaped losses.

Lemma 5 (Li et al. (2022); Kleinberg et al. (2023)) *Let $\ell : [0,1] \to \mathbb{R}$ be a 1-Lipschitz convex function. There exists a distribution φ on $[0,1]$ and a constant $C = (\ell(0) + \ell(1) - 1)/2$ such that*

$$\ell(p) = \mathbb{E}_{v\sim\varphi}|p - v| + C \quad \text{for every } p \in [0,1].$$

Proof Since ℓ is convex and 1-Lipschitz, its subgradient $\nabla\ell(p) \in [-1,1]$ is a non-decreasing function of $p \in [0,1]$. We can additionally assume without loss of generality that $\nabla\ell$ is right continuous on $[0,1)$ because we can replace $\nabla\ell(p)$ by $\lim_{p'\to p^+}\nabla\ell(p')$ while keeping $\nabla\ell(p)$ as a subgradient of ℓ at p. We can also assume without loss of generality that $\nabla\ell(1) = 1$.

We let φ be the distribution on $[0,1]$ whose CDF is $(\nabla\ell(p)+1)/2$. For every $p, v \in [0,1]$,

$$|p - v| = \int_0^1 \mathbb{I}[v \le t < p]\mathrm{d}t + \int_0^1 \mathbb{I}[p \le t < v]\mathrm{d}t.$$

Therefore,

$$\begin{aligned}
\mathbb{E}_{v\sim\varphi}|p-v| &= \int_0^1 \mathbb{E}_{v\sim\varphi}\mathbb{I}[v \le t < p]\mathrm{d}t + \int_0^1 \mathbb{E}_{v\sim\varphi}\mathbb{I}[p \le t < v]\mathrm{d}t \\
&= \int_0^p \Pr_{v\sim\varphi}[v \le t]\mathrm{d}t + \int_p^1 \Pr_{v\sim\varphi}[v > t]\mathrm{d}t \\
&= \int_0^p \frac{\nabla\ell(t)+1}{2}\mathrm{d}t + \int_p^1 \frac{1-\nabla\ell(t)}{2}\mathrm{d}t \\
&= (\ell(p)-\ell(0)+p)/2 + ((1-p)-\ell(1)+\ell(p))/2 \\
&= \ell(p) + (1-\ell(0)-\ell(1))/2 \\
&= \ell(p) - C.
\end{aligned}$$

■

As we discuss in Section 1.2 and formally prove in Appendix D, the minimax theorem allows us to consider the following order-reversed game in lieu of the original swap regret minimization problem (Problem 1):

Problem 6 (Order-reversed online swap regret minimization) *Let T be a positive integer and B be a finite subset of $[0,1]$. We study the following sequential prediction problem. For round $t = 1,\dots,T$,*

1. *adversary reveals a distribution π_t of 1-Lipschitz convex losses $\ell : [0,1] \to \mathbb{R}$;*
2. *predictor chooses $p_t \in B \subseteq [0,1]$;*
3. *loss $\ell_t \sim \pi_t$ is sampled and revealed, and predictor incurs loss $\ell_t(p_t)$.*

The predictor's goal is to minimize the expected swap regret $\mathbb{E}[\mathsf{SR}(p_{1:T},\ell_{1:T})]$ defined in (1).

3. Non-constructive Proof

We show a prediction strategy that achieves $O(\sqrt{T}\log T)$ expected swap regret for the order-reversed game in Problem 6 (Lemma 10). By the minimax argument we formally describe in Appendix D, this implies the same expected swap regret guarantee for Problem 1, thus proving a non-constructive version of Theorem 2 that removes the running time guarantee.

3.1. Width of a Distribution

We introduce the notion of *width* for distributions on the unit interval $[0,1]$ in Lemma 8 below. This notion plays a central role in our multi-scale binning idea for solving Problem 6.

Definition 7 *Let φ be an arbitrary distribution on $[0,1]$. For $z \in [0,1]$, we say $q \in [0,1]$ is a z-quantile of φ if*

$$\Pr_{v\sim\varphi}[v < q] \le z \quad \text{and} \quad \Pr_{v\sim\varphi}[v \le q] \ge z.$$

We use $Q_\varphi(z) \subseteq [0,1]$ to denote the set of all z-quantiles of φ.

Intuitively, if we eventually choose bins of width w, there are two things we want to balance. First is the width w itself, since it contributes directly to the rounding error. The second is the reciprocal of the probability mass in a median-centered interval of this width, since this provides a rough estimate of the number of bins we expect to play (and therefore contributes to the sampling error). The following definition of width balances these two quantities.

Lemma 8 *Let φ be a distribution on $[0,1]$ and let $\gamma \in (0,1]$ be an arbitrary parameter. There exists a unique value $w \in [\gamma, 1]$ such that*

$$w \in Q_\varphi\left(\frac{1}{2} + \frac{\gamma}{2w}\right) - Q_\varphi\left(\frac{1}{2} - \frac{\gamma}{2w}\right).$$

Here, the difference $Q_1 - Q_2$ between two sets Q_1 and Q_2 is defined as the set $\{q_1 - q_2 \,|\, q_1 \in Q_1, q_2 \in Q_2\}$. That is, there exist $\alpha, \beta \in [0,1]$ such that

$$\alpha \in Q_\varphi\left(\frac{1}{2} - \frac{\gamma}{2w}\right), \quad \beta \in Q_\varphi\left(\frac{1}{2} + \frac{\gamma}{2w}\right), \quad \beta - \alpha = w.$$

*We say w is the γ-*width *of φ.*

To prove Lemma 8, we need the following lemma that follows from basic properties of φ as a probability measure.

Lemma 9 *Let φ be a distribution on $[0,1]$. For every $z \in [0,1]$, $Q_\varphi(z) \subseteq [0,1]$ is a non-empty closed interval or a single point. Moreover,*

$$\max Q_\varphi(z) = \begin{cases} 1, & \text{if } z = 1, \\ \inf_{z' \in (z,1]} \min Q_\varphi(z'), & \text{if } z \in [0,1); \end{cases}$$
$$\min Q_\varphi(z) = \begin{cases} 0, & \text{if } z = 0, \\ \sup_{z' \in [0,z)} \max Q_\varphi(z'), & \text{if } z \in (0,1]. \end{cases}$$

Proof [Proof of Lemma 8] For every $w \in [\gamma, 1]$, define

$$S(w) := Q_\varphi\left(\frac{1}{2} + \frac{\gamma}{2w}\right) - Q_\varphi\left(\frac{1}{2} - \frac{\gamma}{2w}\right).$$

By Lemma 9, for every $w \in [\gamma, 1]$, $S(w)$ is a non-empty closed interval or a single point, and $S(w) \subseteq [0,1]$. Our goal is to show that there exists a unique $w^* \in [\gamma, 1]$ such that $w^* \in S(w^*)$. By Lemma 9,

$$\max S(w) = \begin{cases} 1, & \text{if } w = \gamma, \\ \inf_{w' \in [\gamma, w)} \min S(w'), & \text{if } w \in (\gamma, 1]; \end{cases} \tag{9}$$
$$\min S(w) = \sup_{w' \in (w,1]} \max S(w'), \quad \text{if } w \in [\gamma, 1). \tag{10}$$

Define $W := \{w \in [\gamma, 1] \,|\, w \le \max S(w)\}$. By (9), we have $\max S(\gamma) = 1$, so $\gamma \in W$. In particular, W is a non-empty set. Let $w^* := \sup W$. We show that $w^* \in S(w^*)$.

We first show $w^* \le \max S(w^*)$. This holds trivially if $w^* = \gamma$. When $w^* > \gamma$, by (9), for every $w \in (\gamma, w^*) \cap W$, we have $w \le \max S(w) = \inf_{w' \in [\gamma, w)} \min S(w')$. Taking the limit $w \to w^*$, we get $w^* \le \inf_{w' \in [\gamma, w^*)} \min S(w') = \max S(w^*)$.

Now we show $w^* \ge \min S(w^*)$ This holds trivially if $w^* = 1$. When $w^* < 1$, by the definition of $w^* = \sup W$, for every $w \in (w^*, 1]$, we have $w > \max S(w)$. Taking the limit $w \to w^*$, we have $w^* \ge \lim_{w \to (w^*)^+} \max S(w) = \sup_{w \in (w^*, 1]} \max S(w) = \min S(w^*)$ by (10).

We have proved that $\min S(w^*) \le w^* \le \max S(w^*)$, which implies that $w^* \in S(w^*)$. To show that w^* is the unique value with this property, we note that for every $w < w^*$, we have

$$w < w^* \le \max S(w^*) = \inf_{w' < w^*} \min S(w') \le \min S(w),$$

which means that $w \notin S(w)$. Similarly, for every $w > w^*$, we also have $w \notin S(w)$. ■

3.2. Truthful Predictor with Multi-scale Binning

We are now ready to describe our prediction strategy that achieves $O(\sqrt{T} \log T)$ expected swap regret for Problem 6. In fact, we prove a stronger result: given any $\delta \in (0, 1/T]$ as input, our strategy guarantees $O\big(\sqrt{(T \log T) \log(1/\delta)}\big)$ swap regret with probability at least $1 - \delta$.

Let $\gamma \in (0, 1]$ be a parameter we determine later. We define R as the set of bin scales we consider:

$$R := \{\gamma, 2\gamma, 4\gamma, 8\gamma, \ldots\} \cap [0, 1]. \tag{11}$$

We will always ensure $\gamma = \Omega(1/\sqrt{T})$, which implies $|R| = O(\log T)$. For every $r \in R$, we define B_r as the bins at scale r:

$$B_r := \{0, r, 2r, 3r, \ldots\} \cap [0, 1]. \tag{12}$$

Clearly, $|B_r| = O(1/r)$. We define

$$B := \bigcup_{r \in R} B_r \quad \text{and} \quad \Theta := \{(r, b) : r \in R, b \in B_r\}. \tag{13}$$

We have $|B| \le |\Theta| = \sum_{r \in R} |B_r| = O(1/\gamma) = O(\sqrt{T})$. We use the following truthful predictor to achieve the $O\big(\sqrt{(T \log T) \log(1/\delta)}\big)$ swap regret bound for Problem 6.

Truthful Predictor with Multi-scale Binning. We choose $\gamma := \sqrt{(\ln T)(\ln(1/\delta))/T}$. If $\gamma > 1$, the trivial swap regret upper bound T already implies the desired $O\big(\sqrt{(T \log T) \log(1/\delta)}\big)$ bound. We thus assume without loss of generality that $\gamma \le 1$. In each round $t = 1, \ldots, T$, we take the following steps:

1. We receive a distribution π_t from the adversary. Here π_t is a distribution of 1-Lipschitz convex loss functions $\ell : [0, 1] \to \mathbb{R}$. By Lemma 5, each ℓ corresponds to a distribution φ on $[0, 1]$. Thus we can equivalently view π_t as a meta-distribution of distributions φ on $[0, 1]$. We define distribution $\bar{\varphi}_t$ on $[0, 1]$ as the mixture of distributions φ drawn from π_t.

2. We define $w_t \in [\gamma, 1]$ to be the γ-width of $\bar{\varphi}_t$ (Lemma 8). This implies the existence of $\alpha_t, \beta_t \in [0, 1]$ such that

$$\alpha_t \in Q_{\bar{\varphi}_t}\left(\frac{1}{2} - \frac{\gamma}{2w_t}\right), \quad \beta_t \in Q_{\bar{\varphi}_t}\left(\frac{1}{2} + \frac{\gamma}{2w_t}\right), \quad \beta_t - \alpha_t = w_t.$$

3. We let r_t be the unique value in R such that $w_t \in [r_t, 2r_t)$. Since $\beta_t - \alpha_t = w_t \geq r_t$, there exists $b_t \in B_{r_t} \subseteq B$ such that $b_t \in [\alpha_t, \beta_t]$. We output $p_t = b_t$.

4. We observe loss ℓ_t drawn from π_t, and let φ_t be the distribution on $[0, 1]$ corresponding to ℓ_t by Lemma 5.

Lemma 10 *For every integer $T \geq 2$ and every $\delta \in (0, 1/T]$, the truthful predictor above achieves $O\big(\sqrt{(T \log T) \log(1/\delta)}\big)$ swap regret with probability at least $1 - \delta$ for Problem 6. In particular, setting $\delta = 1/T$, we get an $O(\sqrt{T} \log T)$ expected swap regret guarantee.*

To prove Lemma 10, we need the following basic lemma about the median of a random variable as the minimizer of the expected distance.

Lemma 11 *Let φ be a distribution of $v \in [0, 1]$, and let $s \in Q_\varphi(1/2)$ be a median of φ. For every $b \in [0, 1]$, we have*

$$0 \leq \mathbb{E}_{v\sim\varphi}|b - v| - \mathbb{E}_{v\sim\varphi}|s - v| \leq |b - s|(1 - 2\Pr_{v\sim\varphi}[s < b \leq v \text{ or } v \leq b < s]).$$

Proof [Proof of Lemma 11] We assume without loss of generality that $b > s$. The other case $b < s$ can be handled similarly by symmetry. We have

$$|b - v| - |s - v| \leq \begin{cases} b - s, & \text{if } v < b; \\ s - b, & \text{if } v \geq b. \end{cases}$$

Therefore,

$$\mathbb{E}_{v\sim\varphi}[|b - v| - |s - v|] \leq (b - s)(\Pr_{v\sim\varphi}[v < b] - \Pr_{v\sim\varphi}[v \geq b]) = (b - s)(1 - 2\Pr_{v\sim\varphi}[v \geq b]).$$

We also have

$$|b - v| - |s - v| \geq \begin{cases} b - s, & \text{if } v \leq s; \\ s - b, & \text{if } v > s. \end{cases}$$

Therefore,

$$\mathbb{E}_{v\sim\varphi}[|b - v| - |s - v|] \geq (b - s)(\Pr_{v\sim\varphi}[v \leq s] - \Pr_{v\sim\varphi}[v > s]) \geq (b - s)(1/2 - 1/2) = 0.$$

■

Proof [Proof of Lemma 10] Recall $\gamma = \sqrt{(\ln T)(\ln(1/\delta))/T}$. If $\gamma > 1$, we can simply use the trivial swap regret upper bound T to get $\mathsf{SR}(p_{1:T}; \ell_{1:T}) \leq T = O\big(\sqrt{(T \log T) \log(1/\delta)}\big)$. We thus assume $\gamma \leq 1$ from now on.

For every pair $(r, b) \in \Theta$, we let $T_{r,b}$ denote the set of time steps $t \in \{1, \ldots, T\}$ such that $(r_t, b_t) = (r, b)$. For every $t \in T_{r,b}$, we have

$$\Pr_{v\sim\bar{\varphi}_t}[v > b + 2r] \leq \Pr_{v\sim\bar{\varphi}_t}[v > \beta_t] \leq \frac{1}{2} - \frac{\gamma}{2w_t} \leq \frac{1}{2} - \frac{\gamma}{4r},$$
$$\Pr_{v\sim\bar{\varphi}_t}[v < b - 2r] \leq \Pr_{v\sim\bar{\varphi}_t}[v < \alpha_t] \leq \frac{1}{2} - \frac{\gamma}{2w_t} \leq \frac{1}{2} - \frac{\gamma}{4r},$$

$$\Pr_{v\sim\bar{\varphi}_t}[v \geq b] \geq \Pr_{v\sim\bar{\varphi}_t}[v \geq \beta_t] \geq \frac{1}{2} - \frac{\gamma}{2w_t} \geq \frac{1}{2} - \frac{\gamma}{2r},$$
$$\Pr_{v\sim\bar{\varphi}_t}[v \leq b] \geq \Pr_{v\sim\bar{\varphi}_t}[v \leq \alpha_t] \geq \frac{1}{2} - \frac{\gamma}{2w_t} \geq \frac{1}{2} - \frac{\gamma}{2r}.$$

Let $\hat{\varphi}_{r,b}$ be the uniform mixture of the distributions φ_t for $t \in T_{r,b}$. Note that each φ_t, as a random variable, has mean $\bar{\varphi}_t$. Thus by standard martingale concentration bounds (Azuma's inequality, see Lemma 16), with probability at least $1 - \delta^4/100$,

$$\Pr_{v\sim\hat{\varphi}_{r,b}}[v > b + 2r] \leq \frac{1}{2} - \frac{\gamma}{4r} + O\left(\sqrt{\frac{\log(1/\delta)}{|T_{r,b}|}}\right), \tag{14}$$

$$\Pr_{v\sim\hat{\varphi}_{r,b}}[v < b - 2r] \leq \frac{1}{2} - \frac{\gamma}{4r} + O\left(\sqrt{\frac{\log(1/\delta)}{|T_{r,b}|}}\right), \tag{15}$$

$$\Pr_{v\sim\hat{\varphi}_{r,b}}[v \geq b] \geq \frac{1}{2} - \frac{\gamma}{2r} - O\left(\sqrt{\frac{\log(1/\delta)}{|T_{r,b}|}}\right), \tag{16}$$

$$\Pr_{v\sim\hat{\varphi}_{r,b}}[v \leq b] \geq \frac{1}{2} - \frac{\gamma}{2r} - O\left(\sqrt{\frac{\log(1/\delta)}{|T_{r,b}|}}\right). \tag{17}$$

By the union bound, with probability at least $1-\delta$, these inequalities hold for all pairs $(r, b) \in \Theta$. It remains to show that whenever these inequalities hold, the swap regret is $O\big(\sqrt{(T\log T)\log(1/\delta)}\big)$. The following equation follows from the definition of $\gamma := \sqrt{(\ln T)(\ln(1/\delta))/T}$:

$$\ln(1/\delta) = T\gamma^2/\ln T. \tag{18}$$

For every $(r, b) \in \Theta$, arbitrarily choose $s_{r,b} \in Q_{\hat{\varphi}_{r,b}}(1/2)$. We can decompose the swap regret as follows.

$$\begin{aligned}
\mathsf{SR}(p_{1:T}, \ell_{1:T}) &= \sum_{t=1}^{T} \ell_t(p_t) - \inf_{\sigma:[0,1]\to[0,1]} \sum_{t=1}^{T} \ell_t(\sigma(p_t)) \\
&= \sum_{t=1}^{T} \ell_t(p_t) - \sum_{b\in B} \inf_{s\in[0,1]} \sum_{t=1}^{T} \mathbb{I}[p_t = b]\ell_t(s) \\
&\leq \sum_{t=1}^{T} \ell_t(p_t) - \sum_{(r,b)\in\Theta} \inf_{s\in[0,1]} \sum_{t\in T_{r,b}} \ell_t(s) \\
&= \sum_{(r,b)\in\Theta} \sum_{t\in T_{r,b}} \ell_t(p_t) - \sum_{(r,b)\in\Theta} \inf_{s\in[0,1]} \sum_{t\in T_{r,b}} \ell_t(s) \\
&= \sum_{(r,b)\in\Theta} \sum_{t\in T_{r,b}} \mathbb{E}_{v\sim\varphi_t}|b - v| - \sum_{(r,b)\in\Theta} \inf_{s\in[0,1]} \sum_{t\in T_{r,b}} \mathbb{E}_{v\sim\varphi_t}|s - v| \\
&= \sum_{(r,b)\in\Theta} \sum_{t\in T_{r,b}} \mathbb{E}_{v\sim\varphi_t}|b - v| - \sum_{(r,b)\in\Theta} \sum_{t\in T_{r,b}} \mathbb{E}_{v\sim\varphi_t}|s_{r,b} - v| \\
&= \sum_{(r,b)\in\Theta} |T_{r,b}| \cdot \mathbb{E}_{v\sim\hat{\varphi}_{r,b}}[|b - v| - |s_{r,b} - v|].
\end{aligned} \tag{19}$$

Let $C > 0$ be a sufficiently large absolute constant. We say a pair $(r, b) \in \Theta$ is *heavy* if $|T_{r,b}| \geq CTr^2/\ln T$. Otherwise, we say (r, b) is *light*. For every fixed heavy pair (r, b), we have

$$\sqrt{\frac{\ln(1/\delta)}{|T_{r,b}|}} \leq \sqrt{\frac{(\ln(1/\delta))(\ln T)}{Cr^2 T}} = \frac{\gamma}{r\sqrt{C}}. \tag{20}$$

Therefore, when C is sufficiently large, (14) and (15) imply

$$\Pr_{v\sim\hat{\varphi}_{r,b}}[v > b + 2r] \leq 1/2, \qquad \Pr_{v\sim\hat{\varphi}_{r,b}}[v < b - 2r] \leq 1/2,$$

so there exists $s_{r,b} \in Q_{\hat{\varphi}_{r,b}}(1/2)$ such that $|s_{r,b} - b| \leq 2r$. Similarly, (16) and (17) imply

$$\Pr_{v\sim\hat{\varphi}_{r,b}}[v \geq b] \geq 1/2 - O\left(\frac{\gamma}{r}\right), \qquad \Pr_{v\sim\hat{\varphi}_{r,b}}[v \leq b] \geq 1/2 - O\left(\frac{\gamma}{r}\right),$$

so by Lemma 11,

$$\mathbb{E}_{v\sim\hat{\varphi}_{r,b}}[|v-b| - |v - s_{r,b}|] \leq 2r \cdot O\left(\frac{\gamma}{r}\right) = O(\gamma). \tag{21}$$

For every light pair (r, b), we have the reversed version of inequality (20):

$$\sqrt{\frac{\ln(1/\delta)}{|T_{r,b}|}} \geq \frac{\gamma}{r\sqrt{C}}.$$

Therefore, (16) and (17) imply

$$\Pr_{v\sim\hat{\varphi}_{r,b}}[v \geq b] \geq 1/2 - O\left(\sqrt{\frac{\log(1/\delta)}{|T_{r,b}|}}\right), \qquad \Pr_{v\sim\hat{\varphi}_{r,b}}[v \leq b] \geq 1/2 - O\left(\sqrt{\frac{\log(1/\delta)}{|T_{r,b}|}}\right).$$

By Lemma 11 and the trivial bound $|b - s_{r,b}| \leq 1$ for an arbitrary $s_{r,b} \in Q_{\hat{\varphi}_{r,b}}(1/2)$, we have

$$\mathbb{E}_{v\sim\hat{\varphi}_{r,b}}[|v-b| - |v - s_{r,b}|] \leq O\left(\sqrt{\frac{\log(1/\delta)}{|T_{r,b}|}}\right).$$

Therefore,

$$\begin{aligned} |T_{r,b}| \cdot \mathbb{E}_{v\sim\hat{\varphi}_{r,b}}[|v-b| - |v - s_{r,b}|] &= O\left(\sqrt{|T_{r,b}|} \cdot \sqrt{\log(1/\delta)}\right) \\ &= O\left(\sqrt{Tr^2/\log T} \cdot \sqrt{T\gamma^2/\log T}\right) && \text{(by (18))} \\ &= O\left(Tr\gamma/\log T\right). && (22) \end{aligned}$$

Combining (21) and (22), we have

$$\begin{aligned} &\mathsf{SR}(p_{1:T}; \ell_{1:T}) \\ \leq &\sum_{(r,b)\in\Theta} |T_{r,b}| \cdot \mathbb{E}_{v\sim\hat{\varphi}_{r,b}}[|v-b| - |v - s_{r,b}|] \end{aligned}$$

$$
\begin{aligned}
&= \sum_{\text{heavy } (r,b)} |T_{r,b}| \cdot \mathbb{E}_{v \sim \hat{\varphi}_{r,b}}[|v-b| - |v - s_{r,b}|] + \sum_{\text{light } (r,b)} |T_{r,b}| \cdot \mathbb{E}_{v \sim \hat{\varphi}_{r,b}}[|v-b| - |v - s_{r,b}|] \\
&\leq O(\gamma) \cdot \sum_{\text{heavy } (r,b)} |T_{r,b}| + \sum_{\text{light } (r,b)} O\left(Tr\gamma / \log T\right) \\
&\leq O(\gamma) \cdot T + \sum_{r \in R} O\left(T\gamma / \log T\right) && \text{(because } |B_r| = O(1/r)\text{)} \\
&= O(T\gamma) && \text{(because } |R| = O(\log T)\text{)} \\
&= O\left(\sqrt{(T \log T) \log(1/\delta)}\right).
\end{aligned}
$$

■

References

Jacob Abernethy and Shie Mannor. Does an efficient calibrated forecasting strategy exist? In *Proceedings of the 24th Annual Conference on Learning Theory*, pages 809–812. JMLR Workshop and Conference Proceedings, 2011.

Jacob Abernethy, Peter L. Bartlett, and Elad Hazan. Blackwell approachability and no-regret learning are equivalent. In Sham M. Kakade and Ulrike von Luxburg, editors, *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *Proceedings of Machine Learning Research*, pages 27–46, Budapest, Hungary, 09–11 Jun 2011. PMLR. URL https://proceedings.mlr.press/v19/abernethy11b.html.

Kazuoki Azuma. Weighted sums of certain dependent random variables. *Tohoku Mathematical Journal, Second Series*, 19(3):357–367, 1967. doi: 10.2748/tmj/1178243286.

Avrim Blum and Yishay Mansour. From external to internal regret. *Journal of Machine Learning Research*, 8(47):1307–1324, 2007. URL http://jmlr.org/papers/v8/blum07a.html.

Mark Braverman, Jieming Mao, Jon Schneider, and Matt Weinberg. Selling to a no-regret buyer. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 523–538, 2018.

William Brown, Jon Schneider, and Kiran Vodrahalli. Is learning in games good for the learners? *Advances in Neural Information Processing Systems*, 36:54228–54249, 2023.

Linda Cai, S Matthew Weinberg, Evan Wildenhain, and Shirley Zhang. Selling to multiple no-regret buyers. In *International Conference on Web and Internet Economics*, pages 113–129. Springer, 2023.

Modibo K Camara, Jason D Hartline, and Aleck Johnsen. Mechanisms for a no-regret agent: Beyond the common prior. In *2020 ieee 61st annual symposium on foundations of computer science (focs)*, pages 259–270. IEEE, 2020.

Liyu Chen, Haipeng Luo, and Chen-Yu Wei. Impossible tuning made possible: A new expert algorithm and its applications. In Mikhail Belkin and Samory Kpotufe, editors, *Proceedings of Thirty Fourth Conference on Learning Theory*, volume 134 of *Proceedings of Machine Learning Research*, pages 1216–1259. PMLR, 15–19 Aug 2021. URL https://proceedings.mlr.press/v134/chen21f.html.

Yuval Dagan, Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich. From external to swap regret 2.0: An efficient reduction for large action spaces. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, page 1216–1222, New York, NY, USA, 2024a. Association for Computing Machinery. ISBN 9798400703836. doi: 10.1145/3618260.3649681. URL https://doi.org/10.1145/3618260.3649681.

Yuval Dagan, Constantinos Daskalakis, Maxwell Fishelson, Noah Golowich, Robert Kleinberg, and Princewill Okoroafor. Improved bounds for calibration via stronger sign preservation games. *arXiv preprint arXiv:2406.13668*, 2024b.

Yuan Deng, Jon Schneider, and Balasubramanian Sivan. Prior-free dynamic auctions with low regret buyers. *Advances in Neural Information Processing Systems*, 32, 2019a.

Yuan Deng, Jon Schneider, and Balasubramanian Sivan. Strategizing against no-regret learners. *Advances in neural information processing systems*, 32, 2019b.

Maxwell Fishelson, Noah Golowich, Mehryar Mohri, and Jon Schneider. High-dimensional calibration from swap regret. *arXiv preprint arXiv:2505.21460*, 2025a.

Maxwell Fishelson, Robert Kleinberg, Princewill Okoroafor, Renato Paes Leme, Jon Schneider, and Yifeng Teng. Full swap regret and discretized calibration. In Gautam Kamath and Po-Ling Loh, editors, *Proceedings of The 36th International Conference on Algorithmic Learning Theory*, volume 272 of *Proceedings of Machine Learning Research*, pages 444–480. PMLR, 24–27 Feb 2025b. URL https://proceedings.mlr.press/v272/fishelson25a.html.

Dean P Foster and Rakesh V Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21(1-2):40–55, 1997.

Dean P Foster and Rakesh V Vohra. Asymptotic calibration. *Biometrika*, 85(2):379–390, 1998.

David A. Freedman. On tail probabilities for martingales. *The Annals of Probability*, 3(1):100–118, 1975. ISSN 00911798, 2168894X. URL http://www.jstor.org/stable/2959268.

Sumegha Garg, Christopher Jung, Omer Reingold, and Aaron Roth. Oracle efficient online multicalibration and omniprediction. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2725–2792. SIAM, 2024.

Parikshit Gopalan, Adam Tauman Kalai, Omer Reingold, Vatsal Sharan, and Udi Wieder. Omnipredictors. In *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, pages 79–1. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2022.

Parikshit Gopalan, Michael Kim, and Omer Reingold. Swap agnostic learning, or characterizing omniprediction via multicalibration. *Advances in Neural Information Processing Systems*, 36: 39936–39956, 2023.

Parikshit Gopalan, Princewill Okoroafor, Prasad Raghavendra, Abhishek Sherry, and Mihir Singhal. Omnipredictors for regression and the approximate rank of convex functions. In Shipra Agrawal and Aaron Roth, editors, *Proceedings of Thirty Seventh Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pages 2027–2070. PMLR, 30 Jun–03 Jul 2024. URL https://proceedings.mlr.press/v247/gopalan24b.html.

Varun Gupta, Christopher Jung, Georgy Noarov, Mallesh M Pai, and Aaron Roth. Online multivalid learning: Means, moments, and prediction intervals. In *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, pages 82–1. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2022.

Nika Haghtalab, Michael Jordan, and Eric Zhao. A unifying perspective on multi-calibration: Game dynamics for multi-objective learning. *Advances in Neural Information Processing Systems*, 36: 72464–72506, 2023a.

Nika Haghtalab, Chara Podimata, and Kunhe Yang. Calibrated stackelberg games: Learning optimal commitments against calibrated agents. *Advances in Neural Information Processing Systems*, 36: 61645–61677, 2023b.

Sergiu Hart and Andreu Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

Sergiu Hart and Andreu Mas-Colell. *Simple adaptive strategies: from regret-matching to uncoupled dynamics*, volume 4. World Scientific, 2013.

Ursula Hébert-Johnson, Michael Kim, Omer Reingold, and Guy Rothblum. Multicalibration: Calibration for the (computationally-identifiable) masses. In *International Conference on Machine Learning*, pages 1939–1948. PMLR, 2018.

Lunjia Hu and Yifan Wu. Predict to minimize swap regret for all payoff-bounded tasks. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 244–263, 2024. doi: 10.1109/FOCS61266.2024.00024.

Lunjia Hu, Haipeng Luo, Spandan Senapati, and Vatsal Sharan. Efficient swap multicalibration of elicitable properties. *arXiv preprint arXiv:2511.04907*, 2025.

Christopher Jung, Changhwa Lee, Mallesh Pai, Aaron Roth, and Rakesh Vohra. Moment multicalibration for uncertainty estimation. In *Conference on Learning Theory*, pages 2634–2678. PMLR, 2021.

Bobby Kleinberg, Renato Paes Leme, Jon Schneider, and Yifeng Teng. U-calibration: Forecasting for an unknown agent. In Gergely Neu and Lorenzo Rosasco, editors, *Proceedings of Thirty Sixth Conference on Learning Theory*, volume 195 of *Proceedings of Machine Learning Research*, pages 5143–5145. PMLR, 12–15 Jul 2023. URL https://proceedings.mlr.press/v195/kleinberg23a.html.

Nicolas S Lambert. Elicitation and evaluation of statistical forecasts. *Preprint*, 3(5):18, 2011.

Daniel Lee, Georgy Noarov, Mallesh Pai, and Aaron Roth. Online minimax multiobjective optimization: Multicalibeating and other applications. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 29051–29063. Curran Associates, Inc., 2022. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/ba942323c447c9bbb9d4b638eadefab9-Paper-Conference.pdf.

Yingkai Li, Jason D. Hartline, Liren Shan, and Yifan Wu. Optimization of scoring rules. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, EC '22, page 988–989, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391504. doi: 10.1145/3490486.3538338. URL https://doi.org/10.1145/3490486.3538338.

Haipeng Luo, Spandan Senapati, and Vatsal Sharan. Improved bounds for swap multicalibration and swap omniprediction. *arXiv preprint arXiv:2505.20885*, 2025a.

Haipeng Luo, Spandan Senapati, and Vatsal Sharan. Simultaneous swap regret minimization via KL-calibration. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025b. URL https://openreview.net/forum?id=Ib4ZXPXpss.

Yishay Mansour, Mehryar Mohri, Jon Schneider, and Balasubramanian Sivan. Strategizing against learners in bayesian games. In *Conference on Learning Theory*, pages 5221–5252. PMLR, 2022.

Georgy Noarov and Aaron Roth. The statistical scope of multicalibration. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 26283–26310. PMLR, 23–29 Jul 2023. URL https://proceedings.mlr.press/v202/noarov23a.html.

Georgy Noarov, Ramya Ramalingam, Aaron Roth, and Stephan Xie. High-dimensional prediction for sequential decision making. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=uRAgIVnAO6.

Kent Harold Osband. *Providing Incentives for Better Cost Forecasting (Prediction, Uncertainty Elicitation)*. University of California, Berkeley, 1985.

Binghui Peng. High dimensional online calibration in polynomial time. *arXiv preprint arXiv:2504.09096*, 2025.

Binghui Peng and Aviad Rubinstein. Fast swap regret minimization and applications to approximate correlated equilibria. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1223–1234, 2024.

Mingda Qiao and Gregory Valiant. Stronger calibration lower bounds via sidestepping. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2021, page 456–466, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450380539. doi: 10.1145/3406325.3451050. URL https://doi.org/10.1145/3406325.3451050.

Aaron Roth. Uncertain: Modern topics in uncertainty estimation. Unpublished lecture notes, 2022.

Ingo Steinwart, Chloé Pasin, Robert Williamson, and Siyu Zhang. Elicitation and identification of properties. In *Conference on Learning Theory*, pages 482–526. PMLR, 2014.

Appendix A. Efficient Algorithm

In the previous section, we proved the *existence* of a randomized prediction strategy achieving $O(\sqrt{T}\log T)$ expected swap regret. We now explicitly construct such a prediction algorithm to obtain the running time guarantee of Theorem 2.

The challenge is that we can no longer rely on the minimax theorem and work with the order-reversed Problem 6 where we see the distribution π_t of loss functions before making a prediction p_t. Instead, we need to directly work with Problem 1, where we compute a distribution ρ_t of randomized predictions p_t based only on information from previous rounds, with the loss function ℓ_t revealed only after we commit to the distribution ρ_t. Inspired by the analysis in the previous section, in each round we output not just a randomized prediction $p_t = b_t \in B \subseteq [0,1]$, but also a scale $r_t \in R$. Here we define R, B_r, B, Θ as in (11)-(13) in the previous section. As before, for every $(r,b) \in \Theta$, we define $T_{r,b} := \{t \in [T] \,|\, (r_t, b_t) = (r,b)\}$, and define $\hat{\varphi}_{r,b}$ as the uniform mixture of φ_t over $t \in T_{r,b}$. Based on our proof of Lemma 10 in Section 3, it suffices to compute a distribution κ_t of $(r_t, b_t) \in \Theta$ in each round such that inequalities (14)-(17) hold simultaneously for all pairs $(r,b) \in \Theta$ with probability at least $1-\delta$.

The goal of satisfying the constraints (14)-(17) can be naturally formulated as an online multi-objective learning task. We apply ideas in the *Adversary-Moves-First (AMF)* framework of Lee et al. (2022) to extend our prediction strategy for the order-reversed game (Problem 6) to the original game (Problem 1). The idea is to satisfy a different mixture (i.e. convex combination) of the constraints at every round, and use an *expert algorithm* to choose the mixtures so that satisfying these mixtures implies satisfying *every* individual constraint in hindsight. Similarly to the work of Hu and Wu (2024); Hu et al. (2025) and inspired by Noarov et al. (2025), we use the expert algorithm MsMwC from Chen et al. (2021) to obtain a stronger guarantee than that of the usual multiplicative weights algorithm. This stronger guarantee is critical for achieving our final swap regret bound $O(\sqrt{T}\log T)$.

We first formulate the constraints (14)-(17) using functions, and show that every convex combination of the constraints can be satisfied by some randomized prediction which we can compute efficiently via linear programming (Lemma 12).

Access to revealed losses. When we say that a loss function ℓ_t is revealed, the efficient algorithm does not need a symbolic description of the entire function. It only uses the probabilities, under the distribution φ_t from Lemma 5, of the events appearing in the functions $h_{r,b,\xi}$ below. These events have thresholds in the finite set

$$G := \{b, b+2r, b-2r : (r,b) \in \Theta\} \cap [0,1].$$

By the proof of Lemma 5, if $F_t(x) := \Pr_{v\sim\varphi_t}[v \le x]$, then $F_t(x) = (\nabla^+\ell_t(x)+1)/2$, where $\nabla^+\ell_t(x)$ denotes the right derivative for $x < 1$ and is set to 1 at $x = 1$. The corresponding left limit, with the convention $\nabla^-\ell_t(0) = -1$, gives $\Pr_{v\sim\varphi_t}[v < x]$. Thus it is enough to query one-sided subgradients of ℓ_t at the finitely many points in G and then form the expectations in Step 5 of the algorithm below. Alternatively, if the learner only has value access to ℓ_t, then values on a grid of mesh η determine a piecewise-linear approximation within η uniformly over $[0,1]$ by Lipschitzness; replacing each ℓ_t by this approximation changes swap regret by at most $2\eta T$, so choosing $\eta = 1/\mathsf{poly}(T)$ preserves the stated rate. For every pair $(r,b) \in \Theta$ and $\xi \in \{-2,-1,+1,+2\}$, we define function $h_{r,b,\xi} : \Theta \times [0,1] \to [-1,1]$ as follows. For every $(r',b') \in \Theta$ and $v \in [0,1]$,

$$
\begin{aligned}
h_{r,b,+2}(r',b',v) &= \mathbb{I}[(r',b')=(r,b)]\left(\mathbb{I}[v > b+2r] - \left(\frac{1}{2} - \frac{\gamma}{4r}\right)\right),\\
h_{r,b,-2}(r',b',v) &= \mathbb{I}[(r',b')=(r,b)]\left(\mathbb{I}[v < b-2r] - \left(\frac{1}{2} - \frac{\gamma}{4r}\right)\right),\\
h_{r,b,+1}(r',b',v) &= \mathbb{I}[(r',b')=(r,b)]\left(\left(\frac{1}{2} - \frac{\gamma}{2r}\right) - \mathbb{I}[v \geq b]\right),\\
h_{r,b,-1}(r',b',v) &= \mathbb{I}[(r',b')=(r,b)]\left(\left(\frac{1}{2} - \frac{\gamma}{2r}\right) - \mathbb{I}[v \leq b]\right).
\end{aligned}
$$

Lemma 12 *Let $\gamma \in (0,1]$ be a parameter and define R, B_r, B, Θ as in (11)-(13). Let $\bar{h} : \Theta \times [0,1] \to [-1,1]$ be a convex combination of the functions $h_{r,b,\xi}$ for $(r,b) \in \Theta, \xi \in \{-2,-1,+1,+2\}$. That is, $\bar{h}$ is given by weights $w_{r,b,\xi} \geq 0$ satisfying $\sum_{r,b,\xi} w_{r,b,\xi} = 1$:*

$$\bar{h} = \sum_{(r,b)\in\Theta, \xi\in\{-2,-1,1,2\}} w_{r,b,\xi} h_{r,b,\xi}.$$

Then there exists a distribution κ of $(r',b') \in \Theta$ such that

$$\mathbb{E}_{(r',b')\sim\kappa}[\bar{h}(r',b',v)] \leq 0 \quad \text{for every } v \in [0,1]. \tag{23}$$

Moreover, given the weights $w_{r,b,\xi}$ as input, the distribution κ can be computed in time $\mathsf{poly}(|\Theta|)$ by solving a linear program.

We prove Lemma 12 using Lemma 13 below and the minimax theorem.

Lemma 13 *Let $\gamma \in (0,1]$ be a parameter and define R, B_r, B, Θ as in (11)-(13). For every distribution φ of $v \in [0,1]$, there exists $(r',b') \in \Theta$ such that*

$$\mathbb{E}_{v\sim\varphi}[h_{r,b,\xi}(r',b',v)] \leq 0 \quad \text{for every } (r,b) \in \Theta \text{ and } \xi \in \{-2,-1,+1,+2\}.$$

We prove Lemma 13 using the truthful predictor we constructed in Section 3.2.

Proof [Proof of Lemma 13] We apply the truthful predictor in Section 3.2. Define α, β, w as in Lemma 8 for φ and γ. Choose $r' \in R$ so that $w \in [r', 2r')$. Choose $b' \in B_{r'}$ so that $b' \in [\alpha, \beta]$. This ensures that

$$\mathbb{E}_{v\sim\varphi}[h_{r',b',\xi}(r',b',v)] \leq 0 \quad \text{for every } \xi \in \{-2,-1,+1,+2\}.$$

The proof is completed by noting that $h_{r,b,\xi}(r',b',v) = 0$ if $(r,b) \neq (r',b')$. ■

Proof [Proof of Lemma 12] We first prove the existence of a distribution κ satisfying (23). Although we require (23) to hold for every $v \in [0,1]$, we only need to consider a finite set $V \subseteq [0,1]$ of v

with size $|V| = O(|\Theta|)$. This is because the functions $h_{r,b,\xi}$ only depend on comparisons between v and a finite set of values $b + 2r, b - 2r$ and b for $(r, b) \in \Theta$.

For an arbitrary set S, we use Δ_S to denote the set of probability distributions on S. By Lemma 13,

$$\max_{\varphi \in \Delta_V} \min_{(r',b') \in \Theta} \mathbb{E}_{v \sim \varphi}[\bar{h}(r', b', v)] \leq 0.$$

By the minimax theorem,

$$\min_{\kappa \in \Delta_\Theta} \max_{v \in V} \mathbb{E}_{(r',b') \sim \kappa}[\bar{h}(r', b', v)] \leq 0.$$

This implies the existence of distribution κ satisfying (23) for every $v \in V$, as desired.

To see that κ can be computed in $\mathsf{poly}(|\Theta|)$ time, we note that for every $v \in V$, (23) is a linear constraint on the probability masses of κ. Thus we can compute κ by solving a linear program with $O(|\Theta|)$ constraints and variables. ■

Expert algorithm MsMwC of Chen et al. (2021). In each round $t = 1, \ldots, T$, our prediction algorithm first computes a convex combination $\bar{h}_t$ of the constraint functions. It then makes a randomized prediction using the distribution κ_t from Lemma 12. The convex combination $\bar{h}_t$ is given by weights $w^{(t)}_{r,b,\xi} \geq 0$ such that $\sum_{r,b,\xi} w^{(t)}_{r,b,\xi} = 1$:

$$\bar{h}_t = \sum_{r,b,\xi} w^{(t)}_{r,b,\xi} h_{r,b,\xi}. \tag{24}$$

We use the MsMwC algorithm of Chen et al. (2021) to compute these weights. The algorithm has a low external regret guarantee (25) in the following game and runs in time $\mathsf{poly}(T, |\Theta|)$: in each round $t = 1, \ldots, T$,

1. The MsMwC algorithm computes $w^{(t)}_{r,b,\xi} \geq 0$ for every $(r, b) \in \Theta$ and $\xi \in \{-2, -1, +1, +2\}$ such that $\sum_{r,b,\xi} w^{(t)}_{r,b,\xi} = 1$;
2. The MsMwC algorithm receives adversarially chosen reward $u^{(t)}_{r,b,\xi} \in [-1, 1]$ for every $(r, b) \in \Theta$ and $\xi \in \{-2, -1, +1, +2\}$.

The MsMwC algorithm guarantees that for every $(r_0, b_0) \in \Theta$ and $\xi_0 \in \{-2, -1, +1, +2\}$,

$$\sum_{t=1}^{T} \sum_{r,b,\xi} w^{(t)}_{r,b,\xi} u^{(t)}_{r,b,\xi} \geq \sum_{i=1}^{T} u^{(t)}_{r_0,b_0,\xi_0} - O\left(\log T + \sqrt{\sum_{t=1}^{T} \left(u^{(t)}_{r_0,b_0,\xi_0}\right)^2 \log T}\right). \tag{25}$$

Prediction Algorithm. We are now ready to describe our efficient prediction algorithm for solving Problem 1. As in Section 3.2, we choose $\gamma := \sqrt{(\ln T)(\ln(1/\delta))/T}$ and assume without loss of generality that $\gamma \leq 1$. Define R, B_r, B, Θ as in (11)-(13). For round $t = 1, \ldots, T$,

1. Compute weights $w^{(t)}_{r,b,\xi}$ using MsMwC and form the convex combination $\bar{h}_t : \Theta \times [0, 1] \to [-1, 1]$ by (24);
2. Compute κ_t from $\bar{h}_t$ as in Lemma 12;

3. Sample $(r_t, b_t) \in \Theta$ from κ_t and output $p_t := b_t$;

4. Observe loss function $\ell_t : [0, 1] \to \mathbb{R}$. Let φ_t denote the distribution corresponding to ℓ_t;

5. For every $(r, b) \in \Theta$ and $\xi \in \{-2, -1, +1, +2\}$, provide reward

$$u_{r,b,\xi}^{(t)} := \mathbb{E}_{(r',b')\sim\kappa_t}\mathbb{E}_{v\sim\varphi_t} h_{r,b,\xi}(r', b', v)$$

to the MsMwC algorithm.

Lemma 14 *Let $T > 2$ be an integer. Given any $\delta \in (0, 1/T]$, the prediction algorithm above guarantees*

$$\mathsf{SR}(p_{1:T}; \ell_{1:T}) = O\left(\sqrt{(T \log T)\log(1/\delta)}\right)$$

with probability at least $1-\delta$ for Problem 1 and runs in $\mathsf{poly}(T)$ time. In particular, setting $\delta = 1/T$, we get an $O(\sqrt{T}\log T)$ expected swap regret upper bound.

Proof [Proof of Lemma 14] The running time guarantee follows from the corresponding guarantees from Lemma 12 and the efficiency of the MsMwC algorithm. We prove the swap regret guarantee below.

The MsMwC algorithm ensures (25), which means that for every $(r, b) \in \Theta$ and $\xi \in \{-2, -1, +1, +2\}$,

$$\sum_{t=1}^{T} \mathbb{E}_{(r',b')\sim\kappa_t}\mathbb{E}_{v\sim\varphi_t}[\bar{h}_t(r', b', v)] \geq \sum_{t=1}^{T} \mathbb{E}_{(r',b')\sim\kappa_t}\mathbb{E}_{v\sim\varphi_t}[h_{r,b,\xi}(r', b', v)] - O\left(\log T + \sqrt{E_{r,b}\log T}\right), \tag{26}$$

where $E_{r,b} := \sum_{t=1}^{T} \Pr_{(r',b')\sim\kappa_t}[(r', b') = (r, b)]$.

By Lemma 12 and our choice of κ_t, for every $t = 1, \ldots, T$,

$$\mathbb{E}_{(r',b')\sim\kappa_t}\mathbb{E}_{v\sim\varphi_t}[\bar{h}_t(r', b', v)] \leq 0.$$

By (26), for every $(r, b) \in \Theta$ and $\xi \in \{-2, -1, +1, +2\}$,

$$\sum_{t=1}^{T} \mathbb{E}_{(r',b')\sim\kappa_t}\mathbb{E}_{v\sim\varphi_t}[h_{r,b,\xi}(r', b', v)] = O\left(\log T + \sqrt{E_{r,b}\log T}\right).$$

Since (r_t, b_t) is drawn from κ_t, by standard Martingale concentration (Freedman's inequality, see Lemma 17), with probability at least $1 - \delta^4$, for every $(r, b) \in \Theta$ and $\xi \in \{-2, -1, +1, +2\}$,

$$\sum_{t=1}^{T} \mathbb{E}_{v\sim\varphi_t} h_{r,b,\xi}(r_t, b_t, v) = O\left(\log(1/\delta) + \sqrt{E_{r,b}\log(1/\delta)}\right), \tag{27}$$

$$||T_{r,b}| - E_{r,b}| = O\left(\sqrt{E_{r,b}\log(1/\delta)}\right). \tag{28}$$

It remains to prove that the two inequalities above imply $\mathsf{SR}(p_{1:T}; \ell_{1:T}) = O\left(\sqrt{(T\log T)\log(1/\delta)}\right)$.

As in the proof of Lemma 10, we define $\hat{\varphi}_{r,b}$ as the uniform mixture of φ_t for $t \in T_{r,b}$. This allows us to decompose the swap regret as in (19). The following equation follows from the definition of $\gamma := \sqrt{(\ln T)(\ln(1/\delta))/T}$:

$$\ln(1/\delta) = T\gamma^2/\ln T. \tag{29}$$

We say a pair $(r, b) \in \Theta$ is *heavy* if $E_{r,b} \geq CTr^2/\ln T$. Otherwise, we say (r, b) is *light*. For heavy pairs, by (29) we have

$$E_{r,b} \geq CTr^2/\ln T \geq CT\gamma^2/\ln T = C\ln(1/\delta).$$

Therefore, when C is sufficiently large, (28) implies $|T_{r,b}| \geq E_{r,b}/2$, and (27) implies

$$\sum_{t=1}^{T} \mathbb{E}_{v\sim\varphi_t} h_{r,b,\xi}(r_t, b_t, v) = O\left(\sqrt{|T_{r,b}|\log(1/\delta)}\right) \quad \text{for every } \xi \in \{-2, -1, +1, +2\}.$$

This means that inequalities (14)-(17) hold, so we have the same bound as (21):

$$\mathbb{E}_{v\sim\hat{\varphi}_{r,b}}[|v-b| - |v-s_{r,b}|] = O(\gamma). \tag{30}$$

For light pairs, (27) implies

$$|T_{r,b}|\left(\frac{1}{2} - \Pr_{v\sim\hat{\varphi}_{r,b}}[v \geq b]\right) \leq \frac{\gamma}{2r} \cdot |T_{r,b}| + O\left(\sqrt{E_{r,b}\log(1/\delta)}\right) + O(\log(1/\delta)). \tag{31}$$

By the definition of (r, b) being a light pair and (29), we have

$$\sqrt{E_{r,b}\ln(1/\delta)} \leq \sqrt{CTr^2/\ln T} \cdot \sqrt{T\gamma^2/\ln T} = \sqrt{C} \cdot Tr\gamma/\ln T. \tag{32}$$

By (28) and the fact that $r \geq \gamma$, we have

$$|T_{r,b}| \leq E_{r,b} + O\left(\sqrt{E_{r,b}\log(1/\delta)}\right) = O\left(Tr^2/\log T\right) + O(Tr\gamma/\log T) = O\left(Tr^2/\log T\right).$$

Therefore,

$$\frac{\gamma}{2r} \cdot |T_{r,b}| = O(Tr\gamma/\log T). \tag{33}$$

Plugging (32), (33), and (29) into (31), we have

$$|T_{r,b}|\left(\frac{1}{2} - \Pr_{v\sim\hat{\varphi}_{r,b}}[v \geq b]\right) = O(Tr\gamma/\log T) + O(T\gamma^2/\log T) = O(Tr\gamma/\log T).$$

Similarly, we have

$$|T_{r,b}|\left(\frac{1}{2} - \Pr_{v\sim\hat{\varphi}_{r,b}}[v \leq b]\right) = O(Tr\gamma/\log T).$$

Therefore, by Lemma 11,

$$|T_{r,b}| \cdot \mathbb{E}_{v\sim\hat{\varphi}_{r,b}}[|v-b| - |v-s_{r,b}|] = O(Tr\gamma/\log T). \tag{34}$$

Finally, note that (30) and (34) correspond to (21) and (22) in the proof of Lemma 10, respectively. The rest of the proof is the same as the proof of Lemma 10. ∎

Appendix B. Related Work

B.1. Swap Regret Minimization

There has been an extensive literature on swap regret minimization in the online learning literature (Hart and Mas-Colell, 2000; Blum and Mansour, 2007; Hart and Mas-Colell, 2013). A line of work studies the game-theoretic properties of swap regret minimization (Braverman et al., 2018; Deng et al., 2019a,b; Camara et al., 2020; Mansour et al., 2022; Cai et al., 2023; Brown et al., 2023; Haghtalab et al., 2023b). Recently, Dagan et al. (2024a); Peng and Rubinstein (2024) show that, without additional structure on the loss function and action space, swap regret admits lower bounds showing an inherent tradeoff: any algorithm must either incur regret with polynomial dependence on the number of actions (e.g., $\widetilde{O}(\sqrt{nT})$ in the worst case) or else require $\exp(\Omega(1/\varepsilon))$ rounds exponential in the target ε average swap regret. Fishelson et al. (2025b) obtain improved swap regret bound with a convex d-dimensional action space and structured losses, where the assumption of losses includes linearity, smoothness, strongly convex, etc. In particular, Fishelson et al. (2025b) imply an $\widetilde{O}\left(T^{2/3}\right)$ swap regret for convex and Lipschitz losses under single-dimensional action space, which we improve to the near-optimal $\widetilde{O}\left(T^{1/2}\right)$.

B.2. Online Calibration

There is a well-established connection between online calibration and swap regret minimization (Foster and Vohra, 1997, 1998). Calibrated predictions guarantee no swap regret for any best-responding decision makers. Viewed as a swap regret minimization problem, online calibration is more structured than the general setting in the sense that action space is allowed to be arbitrary, but the loss function is selected adversarially from a finite space, indexed by the random state to be predicted. On contrary, our setting considers arbitrarily adversarial loss functions with a convex and Lipschitz structure. The literature on online calibration designs algorithms with low calibration error (Foster and Vohra, 1998; Abernethy and Mannor, 2011; Luo et al., 2025b), which mostly focuses on the ℓ_1 calibration error (Qiao and Valiant, 2021; Dagan et al., 2024b). Hu and Wu (2024) minimizes swap regret for all downstream decision makers for binary states, which equivalently implies two loss functions for the adversary. This binary state results in an improved $\widetilde{O}(\sqrt{T})$ worst-case swap regret for downstream decision makers. Recently, Peng (2025); Fishelson et al. (2025a) show that an exponential $\exp\left(\text{poly}\left(\frac{1}{\epsilon}\right)\right)$ rounds (assuming dimension $d \geq \text{poly}\left(\frac{1}{\epsilon}\right)$) is required for achieving ϵ ℓ_1-calibration error.

B.3. Calibration for Elicitable Properties

Calibration for elicitable properties requires a predicted statistical property is consistent with the same property of the empirical conditional distribution. Section 1.1.1 introduces the connections and distinctions of our swap regret minimization and existing work on calibration for elicitable properties. Jung et al. (2021) study the (multi-)calibration of moments and show it is impossible to calibrated variance and other higher moments. Noarov and Roth (2023) establishes the equivalence between the possibility of (multi-)calibration and elicitability of a statistical property. Section 1.1.1 has introduced the connection of our work to some previous work. The line of previous work mostly focuses on the indentification calibration error in both offline and online setting, e.g., for mean (Hébert-Johnson et al., 2018; Gopalan et al., 2022; Gupta et al., 2022; Haghtalab et al., 2023a; Garg et al., 2024; Luo et al., 2025a; Hu et al., 2025), q-th quantile (Garg et al., 2024; Roth, 2022; Hu et al., 2025).

Distinctions. We note that prior results on calibration for elicitable properties do not apply to our swap regret of scoring functions due to the distinction in definitions. Prior work uses alternative definitions based on *identification functions* of the proper scoring rule, which we refer to as *identification calibration error* for elicitable properties. (Noarov and Roth, 2023; Hu et al., 2025). For any elicitable property Γ, there exists an identification function $V : [0,1] \times Y \to \mathbb{R}$ such that for every distribution q over Y and every $\gamma \in [0,1]$,

$$\mathbb{E}_{y \sim q}[V(\gamma, y)] = 0 \quad \Longleftrightarrow \quad \gamma \in \Gamma(q).$$

Generally, the identification function V is the derivative of a scoring function that elicits Γ: $V(\gamma, y) = \partial_\gamma S(\gamma, y)$ (Osband, 1985; Lambert, 2011; Steinwart et al., 2014). For example, for the median, $V_{\mathsf{median}}(p, y) = 2\mathbb{I}[p \geq y] - 1$; for the mean, $V_{\mathsf{mean}}(p, y) = 2(p - y)$; and for the q-quantile, $V_{q\text{-}\mathsf{quantile}}(p, y) = q - \mathbb{I}[y \leq p]$.

The identification calibration error is defined as the ℓ_r-aggregation of the identification function with results for $r \geq 2$. Given a sequence $(x_t, p_t, y_t)_{t=1}^T$, let $n_p = |\{t \in [T] : p_t = p\}|$. The ℓ_r-calibration error in Noarov and Roth (2023); Hu et al. (2025) is

$$\mathrm{MCal}_r := \sum_{p \in [0,1]} n_p \left| \frac{1}{n_p} \sum_{t=1}^{T} \mathbb{I}[p_t = p] \, V(p_t, y_t) \right|^r . \tag{35}$$

Hu et al. (2025) design an algorithm that achieves $O\left(T^{\frac{1}{r+1}}\right)$ identification calibration error for several assumptions: 1) $r \geq 2$; 2) Lipschitz identification functions; 3) Lipschitz distribution of the states. Notably for 3), Hu et al. (2025) and Noarov and Roth (2023) assume the distribution of the state Y is Lipschitz continuous, while we do not impose any restrictions on the distribution.

Several distinctions arise due to different definitions.

- First, a proper scoring functions have a connection with downstream decision payoff. Given a prediction of a statistical property, the proper scoring function corresponds to the payoff of some decision maker whose decision payoff only depends on the property and best responds to the prediction. Our problem is thus equivalent to minimizing the swap regret of a downstream decision maker.

- Second, the Lipschitzness of the distribution is necessary for median and quantile with identification calibration error in previous work, while our swap regret minimization do not require such an assumption, especially when the algorithm adopts a fixed binning discretization strategy of the prediction space. Intuitively, we take median calibration as an example. When $r = 1$, the identification calibration error of median simplifies to the distance between the empirical quantile of the prediction p and $\frac{1}{2}$:

$$\begin{aligned} \mathrm{MCal}_1^{\mathsf{median}} &= \sum_{p \in [0,1]} n_p \left| 2 \cdot \frac{1}{n_p} \sum_{t:\, p_t = p} \mathbb{I}[y_t \leq p] - 1 \right| \\ &= 2 \sum_{p \in [0,1]} n_p \left| \Pr_{y \sim \tau_p}[y \leq p] - \frac{1}{2} \right|, \end{aligned}$$

i.e., for each prediction value p, we compute the empirical quantile level of p among the associated samples $\{y_t : p_t = p\}$ and compare it to $1/2$.

Lemma 15 shows that if an online prediction algorithm predicts from a fixed set of prediction values (which many previous algorithms do), then there exists a (non-Lipschitz) distribution where the identification calibration error is $\Theta(T)$. The main idea is that, when the true distribution of y_t's is not Lipschitz and concentrates in a fixed bin, the discretization error cannot be bounded for $\mathrm{MCal}_1^{\mathsf{median}}$.

Lemma 15 *Suppose an online prediction algorithm makes predictions from a finite set $B \subseteq [0,1]$. There exists a distribution τ such that when $\forall t, y_t \overset{i.i.d.}{\sim} \tau$, $\mathbb{E}[\mathrm{MCal}_1^{\mathsf{median}}] = \Theta(T)$.*

Proof Let $B \subseteq [0,1]$ be finite, and write its elements in increasing order as

$$b_1 < b_2 < \cdots < b_m.$$

Pick any pair of adjacent values, say $b_i < b_{i+1}$. Define the outcome distribution τ to be uniform on the open interval (b_i, b_{i+1}):

$$y_t \overset{\text{i.i.d.}}{\sim} \tau := \mathrm{Unif}(b_i, b_{i+1}).$$

(Notice τ is non-Lipschitz in the usual sense used in the identification-calibration literature.)

Now consider *any* online prediction algorithm that at each time t outputs $p_t \in B$ (possibly adaptively as a function of the past). For each $p \in B$, let $n_p := |\{t \in [T] : p_t = p\}|$ denote the number of times the algorithm predicts p. Recall that (with the convention that the summand is 0 when $n_p = 0$)

$$\mathrm{MCal}_1^{\mathsf{median}} \;=\; 2 \sum_{p \in B} n_p \left| \frac{1}{n_p} \sum_{t:\, p_t = p} \mathbb{I}[y_t \le p] - \frac{1}{2} \right|.$$

We claim that for every $p \in B$, the indicator $\mathbb{I}[y_t \le p]$ is *deterministic* under our choice of τ:

- If $p \le b_i$, then $y_t > b_i \ge p$ almost surely (since $y_t \in (b_i, b_{i+1})$), so $\mathbb{I}[y_t \le p] = 0$ a.s.
- If $p \ge b_{i+1}$, then $y_t < b_{i+1} \le p$ almost surely, so $\mathbb{I}[y_t \le p] = 1$ a.s.

There is no third case, because b_i and b_{i+1} are *adjacent* in B, hence no element of B lies in (b_i, b_{i+1}).

Therefore, for each fixed $p \in B$, the empirical average inside the absolute value is either 0 (when $p \le b_i$) or 1 (when $p \ge b_{i+1}$), deterministically. In either case,

$$\left| \frac{1}{n_p} \sum_{t:\, p_t = p} \mathbb{I}[y_t \le p] - \frac{1}{2} \right| \;=\; \tfrac{1}{2}, \qquad \text{whenever } n_p > 0.$$

Plugging this into the definition gives

$$\mathrm{MCal}_1^{\mathsf{median}} \;=\; 2 \sum_{p \in B} n_p \cdot \tfrac{1}{2} \;=\; \sum_{p \in B} n_p \;=\; T,$$

since $\sum_{p \in B} n_p = T$ always. Hence the identification calibration error is *exactly* T (not merely in expectation), and in particular

$$\mathbb{E}\big[\mathrm{MCal}_1^{\mathsf{median}}\big] \;=\; T \;=\; \Theta(T).$$

This proves the lemma. ■

Appendix C. Martingale Concentration Bounds

Lemma 16 (Azuma's Inequality (Azuma, 1967)) *Let $X_1, \ldots, X_n \in [-1, 1]$ be a martingale difference sequence. There exists an absolute constant $C > 0$ such that for every $\delta \in (0, 1/2)$, with probability at least $1 - \delta$,*

$$\left| \sum_{i=1}^{n} X_i \right| \le C\sqrt{n \log(1/\delta)}.$$

Lemma 17 (Freedman's Inequality (Freedman, 1975)) *Let $X_1, \ldots, X_n \in [-1, 1]$ be a martingale difference sequence. Let $V_i := \mathbb{E}[X_i^2 | X_1, \ldots, X_{i-1}]$ denote the conditional variance of X_i. There exists an absolute constant $C > 0$ such that for every $\delta \in (0, 1/2)$, with probability at least $1 - \delta$ we have*

$$\left| \sum_{i=1}^{n} X_i \right| \le C\sqrt{\log(n/\delta) \sum_{i=1}^{n} V_i} + C\log(n/\delta).$$

Appendix D. Applying the Minimax Theorem

We use the minimax theorem to complete the proof of the following non-constructive version of Theorem 2 using Lemma 10:

Theorem 18 (Non-constructive version of Theorem 2) *For every positive integer $T \ge 2$, there exists a prediction strategy for Problem 1 that guarantees*

$$\mathbb{E}[\mathsf{SR}(p_{1:T}; \ell_{1:T})] = O(\sqrt{T} \log T).$$

To apply the minimax theorem, we need to carefully specify the strategy spaces of the predictor and the adversary. We note that the swap regret (1) is invariant to adding constants to the loss functions ℓ_t, so we can assume without loss of generality that $\ell_t(0) = 0$ for every $t = 1, \ldots, T$. Let Λ be the family of all 1-Lipschitz convex loss functions $\ell : [0, 1] \to \mathbb{R}$ satisfying $\ell(0) = 0$. A deterministic predictor's strategy $P = (P_1, \ldots, P_T)$ is a sequence of functions where each $P_t : [0, 1]^{t-1} \times \Lambda^{t-1} \to [0, 1]$ decides the prediction $p_t \in [0, 1]$ for the t-th round given the previous $t - 1$ rounds' transcript $(p_1, \ldots, p_{t-1}; \ell_1, \ldots, \ell_{t-1}) \in [0, 1]^{t-1} \times \Lambda^{t-1}$. We use $\mathbb{P}_T$ to denote the set of all such strategies $P = (P_1, \ldots, P_T)$. Similarly, an adversary's strategy is $L = (L_1, \ldots, L_T)$ with $L_t : [0, 1]^{t-1} \times \Lambda^{t-1} \to \Lambda$ for every $t = 1, \ldots, T$. We use $\mathbb{L}_T$ to denote the set of all such L.

Any strategy pair (P, L) uniquely determines the entire game transcript $p_1, \ldots, p_T \in [0, 1]$ and $\ell_1, \ldots, \ell_T \in \Lambda$. This is done by inductively setting

$$p_t = P_t(p_1, \ldots, p_{t-1}; \ell_1, \ldots, \ell_{t-1}) \quad \text{and} \quad \ell_t = L_t(p_1, \ldots, p_{t-1}; \ell_1, \ldots, \ell_{t-1})$$

for $t = 1, \ldots, T$. We define $\mathsf{SR}(P, L) := \mathsf{SR}(p_{1:T}; \ell_{1:T})$. For an arbitrary set S, we use Δ_S to denote the set of probability distributions on S.

Theorem 18 can be equivalently stated as follows:

Lemma 19 *For every positive integer T, there exists a distribution $\mathcal{P}$ on $\mathbb{P}_T$ such that for every $L \in \mathbb{L}_T$,*

$$\mathbb{E}_{P\sim\mathcal{P}}[\mathsf{SR}(P; L)] = O(\sqrt{T} \log T).$$

Equivalently, for every positive integer T,

$$\inf_{\mathcal{P}\in\Delta_{\mathbb{P}_T}} \sup_{L\in\mathbb{L}_T} \mathbb{E}_{P\sim\mathcal{P}}[\mathsf{SR}(P;L)] = O(\sqrt{T}\log T).$$

The minimax theorem requires the strategy spaces of both players to be finite. Towards making the adversary's strategy space finite, we consider a finite ε-net Λ_ε of Λ and fix a discretization map $D_\varepsilon : \Lambda \to \Lambda_\varepsilon$ such that $\|D_\varepsilon(\ell) - \ell\|_\infty \le \varepsilon$ for every $\ell \in \Lambda$. We will use the following simple simulation argument to justify replacing losses by their discretized versions. Given a predictor strategy P, define $\widehat{P}$ to be the strategy that, after seeing a history $(p_1, \dots, p_{t-1}; \ell_1, \dots, \ell_{t-1})$, feeds the discretized history $(p_1, \dots, p_{t-1}; D_\varepsilon(\ell_1), \dots, D_\varepsilon(\ell_{t-1}))$ to P and makes the same prediction as P. Now fix any adversary strategy L and let $(p_t, \ell_t)_{t=1}^T$ be the transcript generated by the interaction of $\widehat{P}$ and L. If we write $\ell_t^\varepsilon := D_\varepsilon(\ell_t)$, then the same prediction sequence $p_{1:T}$ is generated by P on the discretized loss history $\ell_{1:T}^\varepsilon$. Since $\|\ell_t - \ell_t^\varepsilon\|_\infty \le \varepsilon$ for every t, for every fixed swap function σ,

$$\left| \sum_{t=1}^T \ell_t(p_t) - \ell_t(\sigma(p_t)) - \sum_{t=1}^T \ell_t^\varepsilon(p_t) + \ell_t^\varepsilon(\sigma(p_t)) \right| \le 2\varepsilon T.$$

Taking the supremum over σ gives

$$|\mathsf{SR}(\widehat{P};L) - \mathsf{SR}(P;L_\varepsilon)| \le 2\varepsilon T,$$

where $\mathsf{SR}(P;L_\varepsilon)$ denotes the swap regret of P on the discretized transcript above. More formally, for any mixed predictor strategy $\mathcal{P}$ and any original adversary L, the interaction of $\widehat{P}$ with L induces a possibly randomized adversary in the discretized game: after each public history of predictions and discretized losses, this adversary draws the next discretized loss according to the conditional law of $D_\varepsilon(\ell_t)$ in the original interaction. The joint law of $(P, p_{1:T}, \ell_{1:T}^\varepsilon)$ is then the same in the induced discretized game, and because the payoff is linear in the adversary's mixed strategy, a bound against every pure discretized adversary also bounds this induced randomized one. Thus any distribution over strategies P that controls the discretized game gives, by sampling P and running $\widehat{P}$, a strategy for the original game with only an additional $2\varepsilon T$ regret. Since ε can be chosen arbitrarily, it suffices to prove the following lemma:

Lemma 20 *For every $\varepsilon > 0$ and every positive integer T,*

$$\inf_{\mathcal{P}\in\Delta_{\mathbb{P}_T}} \sup_{L\in\mathbb{L}_T} \mathbb{E}_{P\sim\mathcal{P}}[\mathsf{SR}(P;L_\varepsilon)] = O(\sqrt{T}\log T).$$

We prove a stronger version of Lemma 20 by restricting the prediction strategy P to a subset $\mathbb{P}_{T,B}$ of $\mathbb{P}_T$. Here we take a finite subset B of $[0,1]$ and define $\mathbb{P}_{T,B}$ as the set of all prediction strategies $P \in \mathbb{P}_T$ that are restricted to only output predictions $p_t \in B$.

Lemma 21 *For every positive integer T, there exists a finite subset $B \subseteq [0,1]$ such that for every $\varepsilon > 0$,*

$$\inf_{\mathcal{P}\in\Delta_{\mathbb{P}_{T,B}}} \sup_{L\in\mathbb{L}_T} \mathbb{E}_{P\sim\mathcal{P}}[\mathsf{SR}(P;L_\varepsilon)] = O(\sqrt{T}\log T).$$

We are now ready to make the strategy spaces of both players finite. Given $\varepsilon > 0$ and a finite subset $B \subseteq [0,1]$, we define $\mathbb{P}_{T,B,\varepsilon}$ to be the set of function sequences $P = (P_1, \dots, P_T)$ with $P_t : B^{t-1} \times \Lambda_\varepsilon^{t-1} \to B$ for every $t = 1, \dots, T$. Similarly, we define $\mathbb{L}_{T,B,\varepsilon}$ to be the set of function sequences $L = (L_1, \dots, L_T)$ with $L_t : B^{t-1} \times \Lambda_\varepsilon^{t-1} \to \Lambda_\varepsilon$ for every $t = 1, \dots, T$. Lemma 21 is equivalent to the following lemma:

Lemma 22 *For every positive integer T, there exists a finite subset $B \subseteq [0,1]$ such that for every $\varepsilon > 0$,*

$$\inf_{\mathcal{P} \in \Delta_{\mathbb{P}_{T,B,\varepsilon}}} \sup_{L \in \mathbb{L}_{T,B,\varepsilon}} \mathbb{E}_{P \sim \mathcal{P}}[\mathsf{SR}(P; L)] = O(\sqrt{T} \log T).$$

Note that both $\mathbb{P}_{T,B,\varepsilon}$ and $\mathbb{L}_{T,B,\varepsilon}$ are finite sets, so we can apply the minimax theorem on them:

$$\min_{\mathcal{P} \in \Delta_{\mathbb{P}_{T,B,\varepsilon}}} \max_{L \in \mathbb{L}_{T,B,\varepsilon}} \mathbb{E}_{P \sim \mathcal{P}}[\mathsf{SR}(P; L)] = \max_{\mathcal{L} \in \Delta_{\mathbb{L}_{T,B,\varepsilon}}} \min_{P \in \mathbb{P}_{T,B,\varepsilon}} \mathbb{E}_{L \sim \mathcal{L}}[\mathsf{SR}(P; L)].$$

Therefore, Lemma 22 is equivalent to the following lemma:

Lemma 23 *For every positive integer T, there exists a finite subset $B \subseteq [0,1]$ such that for every $\varepsilon > 0$,*

$$\max_{\mathcal{L} \in \Delta_{\mathbb{L}_{T,B,\varepsilon}}} \min_{P \in \mathbb{P}_{T,B,\varepsilon}} \mathbb{E}_{L \sim \mathcal{L}}[\mathsf{SR}(P; L)] = O(\sqrt{T} \log T).$$

Proof [Proof of Theorem 18 using Lemma 10] By our discussion above, it suffices to prove Lemma 23. Given any mixed strategy $\mathcal{L} \in \Delta_{\mathbb{L}_{T,B,\varepsilon}}$ of the adversary, Lemma 10 gives a prediction strategy $P \in \mathbb{P}_{T,B,\varepsilon}$ that achieves $\mathbb{E}_{L \sim \mathcal{L}}[\mathsf{SR}(P; L)] = O(\sqrt{T} \log T)$, completing the proof. ■