Predict to Minimize Swap Regret for All Payoff-Bounded Tasks

Lunjia Hu*
Stanford University
lunjia@stanford.edu

Yifan Wu†
Northwestern University
yifan.wu@u.northwestern.edu

Abstract

A sequence of predictions is calibrated if and only if it induces no swap regret to all downstream decision tasks. We study the Maximum Swap Regret (MSR) of predictions for binary events: the swap regret maximized over all downstream tasks with bounded payoffs. Previously, the best online prediction algorithm for minimizing MSR is obtained by minimizing the K_1 calibration error, which upper bounds MSR up to a constant factor. However, recent work (Qiao and Valiant, 2021) gives an $\Omega(T^{0.528})$ lower bound for the worst-case expected K_1 calibration error incurred by any randomized algorithm in T rounds, presenting a barrier to achieving better rates for MSR. Several relaxations of MSR have been considered to overcome this barrier, via external regret (Kleinberg et al., 2023) and regret bounds depending polynomially on the number of actions in downstream tasks (Noarov et al., 2023; Roth and Shi, 2024). We show that the barrier can be surpassed without any relaxations: we give an efficient randomized prediction algorithm that guarantees $O(\sqrt{T}\log T)$ expected MSR. We also discuss the economic utility of calibration by viewing MSR as a decision-theoretic calibration error metric and study its relationship to existing metrics.

1 Introduction

Given a sequence of predictions indicating, say, the chance of rain for a period of T days, an intuitive way to assess the quality of these predictions is to check for calibration: among the days predicted to have a 60% chance of rain, is the fraction of rainy days indeed 60%? Formally, suppose n_i days receive a prediction of $q_i \in [0,1]$, and among the n_i days, m_i days are actually rainy. Calibration, a notion that originated from the forecasting literature (Dawid, 1982), requires the relationship $m_i = q_i n_i$ to hold for every prediction value q_i. A popular metric for quantifying the calibration error is the K_1 calibration error (a.k.a. Expected Calibration Error, ECE), defined as the sum of $|m_i - q_i n_i|$ over all prediction values q_i that appear in the sequence.

Calibration allows payoff-maximizing decision makers to trust the predictions. Consider a decision maker who needs to choose an action $a \in A$ to maximize their payoff $U(a,\theta)$, a function of the action a (e.g. take an umbrella or not) and the true binary state $\theta \in \{0,1\}$ (e.g. rainy or not). When the true state θ is unknown and only a prediction $p \in [0,1]$ is given, the decision maker can simply trust the prediction and choose the action a that maximizes the expected payoff $\mathbb{E}_{\theta\sim p}[U(a,\theta)]$, *assuming* that $\Pr[\theta = 1]$ is indeed p. When a sequence of predictions is calibrated, this "best response" strategy is optimal for the decision maker to maximize total payoff if no additional information about the true states is provided. Here, the decision making process is decomposed into two

*Supported by Moses Charikar's and Omer Reingold's Simons Investigators awards and the Simons Foundation Collaboration on the Theory of Algorithmic Fairness.

†Supported by NSF CCF-2229162.

separate components: a predictor who aims at making good (e.g. calibrated) predictions without worrying about the decision maker's payoff, and a decision maker who simply best responds to the prediction without needing to gather additional information.

From this decision making perspective, the value of calibration is characterized by the *no swap regret* guarantee (Foster and Vohra, 1997; Kleinberg et al., 2023). That is, any decision maker who best responds to a sequence of calibrated predictions cannot improve their total payoff in retrospect by swapping to a different action $\sigma(a)$ whenever they chose action a. Concretely, for a sequence of actions $\boldsymbol{a} = (a_1, \ldots, a_T)$ chosen by a decision maker with payoff function U, the *swap regret* with respect to the realized trues states $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_T)$ is

$$\textsc{ASwap}_U(\boldsymbol{a}, \boldsymbol{\theta}) := \max_{\sigma: A \to A} \sum_{t \in [T]} \Big[U(\sigma(a_t), \theta_t) - U(a_t, \theta_t) \Big], \tag{1}$$

where A is the set of actions the decision maker can choose from, and the maximum is over all functions that change each action $a \in A$ to another action $\sigma(a) \in A$. A sequence $\boldsymbol{p} = (p_1, \ldots, p_T)$ of predictions is calibrated, if and only if every downstream decision maker, each with their own utility function U and action set A, incurs zero swap regret when choosing their actions $\boldsymbol{a} = (a_1, \ldots, a_T)$ by best responding to $\boldsymbol{p}$. This equivalence between calibration and no swap regret for all decision tasks is observed by Foster and Vohra (1997), and a quantitative version is made explicit by Kleinberg et al. (2023)—the swap regret incurred by any best responding decision maker with a bounded payoff function is upper bounded, up to a constant factor, by the K_1 calibration error.

Swap regret minimization algorithms have been studied extensively in the online learning literature (e.g. Blum and Mansour, 2007; Hart and Mas-Colell, 2013; Peng and Rubinstein, 2023). In game theory, the swap regret is known for its connection to correlated equilibrium. If players in a game all have vanishing swap regret, the empirical distribution of actions converges to correlated equilibria (Foster and Vohra, 1997). Calibration is a powerful tool for swap regret minimization because it allows us to minimize the swap regret not just for one decision task, but simultaneously for *every* decision task. In the *online binary prediction* setting, Foster and Vohra (1998) show a randomized algorithm that makes T sequential predictions $p_1, \ldots, p_T \in [0, 1]$ and guarantees that the expected K_1 calibration error is at most $O(T^{2/3})$, where the prediction p_t at each round t is chosen based on the realized outcomes $\theta_1, \ldots, \theta_{t-1} \in \{0, 1\}$ only in previous rounds. Existence proofs and constructions of such algorithms have been further explored in several subsequent works (Hart, 2022; Foster and Hart, 2021). Consequently, for any downstream decision task, if the decision maker best responds to the predictions, the swap regret grows at this sub-linear rate $O(T^{2/3})$.

The focus of our work is to design online binary prediction algorithms that minimize the Maximum Swap Regret (MSR), which we define as the swap regret maximized over all decision makers, with payoffs bounded in $[0, 1]$, who best respond to our predictions. We restrict the payoffs to be bounded solely for the purpose of normalization: the swap regret scales proportionally if we multiply the payoffs by any positive constant, and it remains the same if any constant is added to the payoffs. Beyond that, we make no additional restrictions on the decision tasks. In particular, we allow each decision task to have an arbitrarily large action space A, and the MSR is the swap regret maximized over all such tasks.

Since the K_1 calibration error is a constant-factor upper bound for MSR, the algorithm by Foster and Vohra (1998) allows us to guarantee $O(T^{2/3})$ MSR in T rounds. This is the best upper-bound guarantee on the MSR known before our work, and it does not match the best known lower bound of $\Omega(\sqrt{T})$.[1] Recently, Qiao and Valiant (2021) prove an $\Omega(T^{0.528})$ lower bound for the K_1 calibration error in online binary prediction. This lower bound presents a barrier to MSR

[1] The $\Omega(\sqrt{T})$ lower bound can be obtained by considering the decision task with two actions ($A = \{0, 1\}$) and the

minimization, showing that new ideas beyond K_1 calibration error minimization are required to achieve, if possible, an $O(\sqrt{T})$ guarantee for MSR.

To overcome this barrier from K_1, several relaxations of MSR have been considered by recent works to achieve $\widetilde{O}(\sqrt{T})$ regret rates in online binary prediction. Kleinberg et al. (2023) show that an $O(\sqrt{T})$ rate can be achieved for the *U-calibration* error, defined as the *external* regret maximized over payoff-bounded tasks, where the σ in (1) is restricted to be a constant function. On the other hand, the work of Noarov et al. (2023); Roth and Shi (2024) shows $\widetilde{O}(|A|\sqrt{T})$ swap regret bounds that depend additionally on the number of actions $|A|$ in the downstream decision task. A natural question remains to be answered: is it indeed necessary to make these relaxations to MSR to outperform the $\Omega(T^{0.528})$ lower bound for the K_1 calibration error?

1.1 Our Contributions

Our main contribution is to show that no relaxation of MSR is needed to overcome the $\Omega(T^{0.528})$ barrier. We give a randomized algorithm that guarantees $O(\sqrt{T}\log T)$ expected MSR in online binary prediction (Theorem 5.1), matching the $\Omega(\sqrt{T})$ lower bound up to a logarithmic factor. In addition, our algorithm is efficient: it runs in $\mathsf{poly}(T)$ time. Thus, our work establishes a much more comprehensive understanding of the optimal rate for MSR, compared to the significant gaps in the current best upper and lower bounds for the following two calibration measures that have recently received significant attention in online binary prediction (see Section 1.2 for more details):

- K_1 calibration error: $O(T^{2/3})$ and $\Omega(T^{0.528})$ (Foster and Vohra, 1998; Qiao and Valiant, 2021);
- Distance to calibration: $O(T^{1/2})$ and $\Omega(T^{1/3})$ (Qiao and Zheng, 2024; Arunachaleswaran et al., 2024).

Our algorithm works in the standard binary prediction setting. In each round $t = 1, \ldots, T$, our algorithm makes a prediction $p_t \in [0, 1]$ and an adversary reveals the true state $\theta_t \in \{0, 1\}$. Both the prediction p_t and the state θ_t can depend on the past history $h_{t-1} = (p_1, \theta_1, \ldots, p_{t-1}, \theta_{t-1})$, but they cannot depend on each other. That is, we can assume without loss of generality that the algorithm chooses p_t and the adversary reveals θ_t simultaneously. This allows our algorithm to leverage the power of randomization: it chooses p_t at random, and the adversary's choice of θ_t cannot depend on the random bits used by our algorithm to generate p_t. Randomization is necessary to guarantee a sub-linear $o(T)$ rate for MSR: without randomization, the adversary can pick $\theta_t = 0$ if the deterministic choice of p_t is above 0.5, and pick $\theta_t = 1$ otherwise. This would lead to at least $T/2$ swap regret for the downstream task with two actions ($A = \{0, 1\}$) and the 0-1 payoff function $U(a, \theta) = \mathbb{I}[a = \theta]$.[2]

While our result is closely related to the main result of Kleinberg et al. (2023), we use substantially different techniques. Kleinberg et al. (2023) achieve an $O(\sqrt{T})$ rate for the *U-calibration error*, defined as the *external* regret maximized over payoff-bounded downstream tasks. The external regret is weaker (i.e. no bigger) than the swap regret because it restricts the function σ in (1) to be a constant function. Focusing on the external regret allows Kleinberg et al. (2023) to reduce general decision tasks to those with only two actions. This reduction leverages the decomposition of general payoffs into *V-shaped payoffs* by Li (r) al. (2022). The prediction algorithm of Kleinberg et al. (2023) ensures that each best responding decision maker takes actions as if they are locally running the classic Hedge algorithm (see Arora et al., 2012). Their external regret guarantee then

0-1 payoff function $U(a, \theta) = \mathbb{I}[a = \theta]$. When the states are independent and unbiased coin flips, the total payoff drops below $T/2 - \Omega(\sqrt{T})$ with constant probability, inducing $\Omega(\sqrt{T})$ swap regret (and $\Omega(\sqrt{T})$ external regret).

[2] We use $\mathbb{I}[\text{statement}]$ to denote the indicator function, which is 1 if the statement is true, and 0 otherwise.

follows directly from the guarantee of the Hedge algorithm. Both the reduction to two actions and the Hedge algorithm are specific to the external regret: they do not directly extend to the stronger notion of swap regret that we focus on.

Although our rate has an additional $\log(T)$ factor compared to the work of Kleinberg et al. (2023), we show it for the much stronger notion of MSR with the external regret in U-calibration replaced by the swap regret (see Example 4.5 where we show $\Omega(T)$ VS $o(1)$ separation between MSR and U-calibration). Our $O(\sqrt{T}\log T)$ rate is independent of the action spaces A in downstream tasks and thus can be viewed as a strengthening of the $\widetilde{O}(|A|\sqrt{T})$ swap regret guarantee by Noarov et al. (2023) and Roth and Shi (2024) for binary states.

We describe our technical ideas behind our main result in Section 1.1.1 below. In addition, we discuss the view of MSR as a calibration error metric in Section 1.1.2, where we present additional results about MSR: efficient algorithms for computing MSR, its connections to other calibration error metrics, and the economic utility of calibration, which justifies the choice of MSR as a calibration error metric.

1.1.1 Online MSR Minimization: Technical Overview

The key idea behind our $O(\sqrt{T}\log T)$ guarantee for MSR comes from the observation that the MSR can often be significantly smaller than the K_1 calibration error, despite their linear relationship in the worst-case. This allows us to bypass the $\Omega(T^{0.528})$ lower bound for the K_1 calibration error. Here we describe a typical example where the MSR is much smaller than the K_1 calibration error. Based on the intuition from this example, we establish a general lemma (Lemma 1.1) which plays a crucial role in our analysis.

Let us discretize the prediction space by restricting our predictor to only produce predictions in a finite set $Q := \{q_1, \dots, q_m\} \subseteq [0,1]$, where $q_i = i/m$ for $i = 1, \dots, m$. We view each q_i as a "bucket", and thus our predictor makes predictions that fall into these buckets. For a sequence of T predictions $\boldsymbol{p} = (p_1, \dots, p_T) \in Q^T$ made by our predictor and the corresponding true states $\boldsymbol{\theta} = (\theta_1, \dots, \theta_T) \in \{0,1\}^T$, let n_i denote the number of predictions in bucket i:

$$n_i := \sum_{t=1}^{T} \mathbb{I}\left[p_t = q_i\right], \tag{2}$$

and let $\widehat{q}_i$ denote the empirical average of the true outcomes corresponding to the n_i predictions:

$$\widehat{q}_i := \frac{1}{n_i} \sum_{t=1}^{T} \theta_t \mathbb{I}\left[p_t = q_i\right]. \tag{3}$$

The K_1 calibration error can be computed as

$$K_1(\boldsymbol{p}, \boldsymbol{\theta}) = \sum_{i=1}^{m} n_i |q_i - \widehat{q}_i|. \tag{4}$$

Assuming $\sqrt{T}$ is an integer, let us choose $m = \sqrt{T}$. Assume for simplicity that each bucket contains the same number of predictions: $n_i = \sqrt{T}$ for every $i = 1, \dots, m$. A typical guarantee one can often obtain using existing online learning techniques is the following bound on the deviation between $\widehat{q}_i$ and q_i:

$$|\widehat{q}_i - q_i| \lesssim 1/\sqrt{n_i} = T^{-1/4}.$$

Thus, let us assume $|\widehat{q}_i - q_i| = T^{-1/4}$ for $i = 1, \dots, m$ for simplicity. Plugging this into (4), we get $K_1(\boldsymbol{p}, \boldsymbol{\theta}) = T^{3/4}$. This is even worse than the $O(T^{2/3})$ upper bound on the K_1 calibration error

guaranteed by the algorithm of Foster and Vohra (1998), mainly because here we discretize the prediction space into $m = \sqrt{T}$ buckets, but the optimal choice would be $m \approx T^{1/3}$.

However, the MSR in this example is much smaller: in fact we have MSR $\leq O(T^{1/2})$. To prove this fact, let us first consider a specific decision task with action space $A = [0, 1]$, where the payoff is given by

$$U(a, \theta) = \mathbb{I}\left[\theta = \mathbb{I}\left[a > 0.5\right]\right] = \begin{cases} \mathbb{I}\left[\theta = 0\right], & \text{if } a \leq 0.5; \\ \mathbb{I}\left[\theta = 1\right], & \text{if } a > 0.5. \end{cases}$$

Here, actions in the interval $[0, 0.5]$ are equivalent in terms of their payoffs, and similarly for actions in $(0.5, 1]$. Therefore, the best response strategy is not unique. One best response strategy which we focus on here is the identity function: $a^*(p) = p$. This induces the actions $\boldsymbol{a} := (a_1, \ldots, a_T) = (p_1, \ldots, p_T)$. We can decompose the swap regret bucket-wise as follows:

$$\begin{aligned}
&\textsc{ASwap}_U(\boldsymbol{a}, \boldsymbol{\theta}) \\
&= \max_{\sigma:[0,1]\to[0,1]} \sum_{t\in[T]} \Big[U(\sigma(a_t), \theta_t) - U(a_t, \theta_t)\Big] \\
&= \max_{\sigma:[0,1]\to[0,1]} \sum_{t\in[T]} \Big[U(\sigma(p_t), \theta_t) - U(p_t, \theta_t)\Big] \\
&= \max_{\sigma:[0,1]\to[0,1]} \sum_{i=1}^{m} \sum_{t=1}^{T} \mathbb{I}\left[p_t = q_i\right] \Big[U(\sigma(q_i), \theta_t) - U(q_i, \theta_t)\Big] \\
&= \max_{\sigma:[0,1]\to[0,1]} \sum_{i=1}^{m} n_i \, \mathbb{E}_{\theta\sim\widehat{q}_i}[U(\sigma(q_i), \theta) - U(q_i, \theta)] \qquad (\theta \in \{0, 1\} \text{ is drawn such that } \Pr[\theta = 1] = \widehat{q}_i) \\
&= \sum_{i=1}^{m} n_i \, \mathbb{E}_{\theta\sim\widehat{q}_i}[U(\widehat{q}_i, \theta) - U(q_i, \theta)]. \qquad (5)
\end{aligned}$$

If q_i and $\widehat{q}_i$ belong to the same half of the interval $[0, 1]$, i.e., $q_i, \widehat{q}_i \in [0, 0.5]$ or $q_i, \widehat{q}_i \in (0.5, 1]$, we clearly have

$$\mathbb{E}_{\theta\sim\widehat{q}_i}[U(\widehat{q}_i, \theta) - U(q_i, \theta)] = 0. \qquad (6)$$

If q_i and $\widehat{q}_i$ belong to different halves of the interval, a simple calculation gives

$$\mathbb{E}_{\theta\sim\widehat{q}_i}[U(\widehat{q}_i, \theta) - U(q_i, \theta)] = \max(\widehat{q}_i, 1 - \widehat{q}_i) - \min(\widehat{q}_i, 1 - \widehat{q}_i) = 2|\widehat{q}_i - 0.5| \leq 2|q_i - \widehat{q}_i|. \qquad (7)$$

Moreover, if q_i and $\widehat{q}_i$ belong to different halves of the interval, by our assumption of $|\widehat{q}_i - q_i| = T^{-1/4}$, we have $|q_i - 0.5| \leq T^{-1/4}$. Plugging (6) and (7) into (5), we get

$$\textsc{ASwap}_U(\boldsymbol{a}, \boldsymbol{\theta}) \leq 2 \sum_{i=1}^{m} n_i |q_i - \widehat{q}_i| \mathbb{I}\left[|q_i - 0.5| \leq T^{-1/4}\right]. \qquad (8)$$

Note that the number of i's satisfying $|q_i - 0.5| \leq T^{-1/4}$ is $O(T^{1/4})$. Thus, by our assumption of $n_i = \sqrt{T}, |q_i - \widehat{q}_i| = T^{-1/4}$, we get $\textsc{ASwap}_U(\boldsymbol{a}, \boldsymbol{\theta}) = O(T^{1/2})$.

While this swap regret bound is for a specific decision task, we can extend it to *every* payoff-bounded decision task and get $\text{MSR}(\boldsymbol{p}, \boldsymbol{\theta}) = O(T^{1/2})$. This follows from a result of Li (r) al. (2022) showing that any payoff-bounded decision task can be expressed as a convex combination of problems with *V-shaped* payoffs (see Section 3.2 for more details). Such decision tasks are all very similar to the one we consider here, and we can similarly obtain an $O(T^{1/2})$ bound for the

swap regret for each of them. This reduction to V-shaped payoffs also played a crucial role in the $O(T^{1/2})$ U-calibration guarantee of Kleinberg et al. (2023).

In our example above, we make the simplifying assumption that the number of predictions in each bucket is the same. However, the $O(T^{1/2})$ bound on the MSR holds without this assumption while only losing a poly-logarithmic factor. That is, we have the following lemma:

Lemma 1.1 (Informal special case of Lemma 5.2 with $\alpha = O(1), \beta = 0, \mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta}) = 0$)**.** *Let T, m be positive integers satisfying $m = \Theta(\sqrt{T})$. Define $Q = \{q_1, \ldots, q_m\} \subseteq [0,1]$ where $q_i = i/m$ for every $i = 1, \ldots, m$. Given a sequence of predictions $\boldsymbol{p} = (p_1, \ldots, p_T) \in Q^T$ and realized states $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_T) \in \{0,1\}^T$, define n_i and $\widehat{q}_i$ as in (2) and (3).*

Assume $|\widehat{q}_i - q_i| \leq O(\frac{1}{\sqrt{n_i}})$ for every $i = 1, \ldots, m$. Then

$$\mathrm{MSR}(\boldsymbol{p}, \boldsymbol{\theta}) \leq \widetilde{O}(T^{1/2}).$$

The intuition behind the lemma can be understood by analyzing the contribution of each bucket to MSR. In the K_1 calibration error, as expressed in (4), the bias $n_i|q_i - \widehat{q}_i|$ in every bucket contributes to the sum, but in our example above, only a minority of the buckets make positive contribution to the MSR, as shown in (8). In general, for any specific decision task with V-shaped payoff, we show that the contribution to MSR from many buckets i is significantly less than the bias $n_i|q_i - \widehat{q}_i|$. This allows us to prove the upper bound on MSR in Lemma 1.1, which would not hold if MSR were replaced by the K_1 calibration error.

Given Lemma 1.1, it remains to show that the guarantee $|\widehat{q}_i - q_i| \leq O(1/\sqrt{n_i})$ can indeed be (approximately) achieved in the online binary prediction setting. We use the result from Noarov et al. (2023) which, as stated, shows an efficient algorithm that gives us a bound only on the maximum of the expectation $\max_i(\mathbf{E}\left[|\widehat{q}_i - q_i|\right] - O(1/\sqrt{n_i}))$, where the expectation is over the randomness in the algorithm. We show that a slight refinement of the analysis in fact gives a bound on the expectation of the maximum $\mathbf{E}\left[\max_i(|\widehat{q}_i - q_i| - O(1/\sqrt{n_i}))\right]$ (Lemma 5.6). As we show in Lemma 5.2 (a generalized version of Lemma 1.1), this bound is sufficient for us to obtain $\mathbb{E}[\mathrm{MSR}(\boldsymbol{p}, \boldsymbol{\theta})] \leq O(\sqrt{T}\log T)$. We also note that a simpler minimax proof, similar to the one by Hart (2022), also allows us to show the existence of a randomized algorithm that approximately guarantees $|\widehat{q}_i - q_i| \leq O(1/\sqrt{n_i})$ and thus, by our Lemma 5.2 again, achieves $\mathbb{E}[\mathrm{MSR}(\boldsymbol{p}, \boldsymbol{\theta})] \leq O(\sqrt{T}\log T)$ (see Appendix A). However, this proof does not come with an explicit construction or any efficiency guarantee.

1.1.2 MSR as a Decision-theoretic Calibration Error

We view MSR as a calibration error metric and study its properties. In Section 3.3, we show MSR can be computed in time polynomial in the size of the prediction space $|Q|$ and the state space $|\Theta|$ by solving a linear program, which holds for general non-binary state space.

In Section 4, we discuss the connection of MSR to other calibration error metrics, including the K_1 and K_2 calibration error, and the smooth calibration error. We consider the average calibration errors across T rounds. The connections show previous results on minimizing K_2 and smCal do not directly apply. We also compare MSR to the U-calibration error UCal in Kleinberg et al. (2023). Note that $\mathrm{UCal} = 0$ is necessary but not sufficient for the predictions to be calibrated.

- K_1 is polynomially related to MSR.

$$\left(\frac{K_1}{T}\right)^2 \leq \frac{\mathrm{MSR}}{T} \leq \frac{2K_1}{T}$$

We give examples where inequalities are asymptotically tight. In fact, the lower bound is achieved by the same example in Section 1.1.1.

- K_2, defined as the sum over squared bias, is polynomially related to MSR.

$$\frac{K_2}{T} \leq \frac{\text{MSR}}{T} \leq 2\sqrt{\frac{K_2}{T}}.$$

We give examples where inequalities are asymptotically tight. There exists an online algorithm that achieves $\widetilde{O}(\sqrt{T})$ K_2 (Roth, 2022).

- The smooth calibration error is not polynomially related to MSR, where we give examples.
- The U-calibration error lowerbounds MSR, but is not polynomially related.

We discuss the economic foundation of MSR in Section 6. Our MSR is a swap regret based error metric. Fixing a decision task, the swap regret is a natural calibration error metric representing the economic utility of calibration. To see this, when the states are stochastically drawn (a.k.a. the batch setting), there exists an *upperbound* on the decision payoff that can be obtained from a prediction - the payoff by best responding to the Bayesian posterior of receiving a prediction. This payoff upperbound is defined as the value of information in the economics literature (Blackwell, 1951). Perfectly calibrated predictions equal the Bayesian posteriors of themselves. For miscalibrated predictions, best responding induces a loss from this payoff upperbound. Swap regret, by swapping the prediction to its Bayesian posterior, measures the loss from the upperbound on the particular decision task. Since calibration is a robust guarantee for all decision tasks, MSR as a calibration error calculates the worst case loss from the upperbound for all decision tasks.

Broader Implications: Decision Payoff vs. Calibration Error As we discuss in Section 6, calibration alone is not related to the payoff on a decision task. Indeed, there can be a discrepancy between the goal of minimizing calibration error and the goal of maximizing decision payoff. Consider the existence proof of the sub-linear K_1 result. Via the minimax theorem, the adversary decides the distribution of the state first, then the algorithm selects a prediction. Qiao and Valiant (2021) show, even if the algorithm knows the distribution of the state, $\widetilde{O}(T^{2/3})$ is the lowerbound of the payoff from reporting truthfully. However, by misreporting the distribution, the algorithm achieves lower K_1 calibration error, which in turn hurts the decision payoff. We provide the minimax proof of $O(\sqrt{T}\log T)$ MSR in Appendix A. In our minimax proof, by reporting truthfully, the algorithm achieves an $O(\sqrt{T}\log T)$ MSR, matching the $\Omega(\sqrt{T})$ lowerbound up to a logarithmic factor. This observation highlights the need of a decision-theoretic foundation to the calibration error metric.

1.2 Related Work

Calibration. While perfect calibration has an intuitive and clear definition, it is a non-trivial and subtle question to meaningfully quantify the calibration error of predictions that are not perfectly calibrated. The K_1 calibration error is one of the most popular calibration measures, but it lacks continuity: slightly perturbing perfectly calibrated predictions can significantly increase the K_1 calibration error. To address this issue, recently Błasiok et al. (2023a) developed a theory of *consistent* calibration measures by introducing the *distance to calibration* as a central notion. This theory has facilitated rigorous explanations of an interesting empirical phenomenon called "calibration out of the box" in deep learning (Błasiok et al., 2023b, 2024). Another way to meaningfully quantify the

calibration error is by measuring the economic utility of the predictions to downstream decision makers. This motivated Kleinberg et al. (2023) to introduce U-calibration and show an optimal $O(\sqrt{T})$ guarantee for it in online binary prediction. Our work can be viewed as a step further towards this direction.

Recent research has made significant progress in proving upper and lower bounds on the optimal rate achievable for both the K_1 calibration error and the distance to calibration in online binary prediction, though significant gaps remain between the current best upper and lower bounds. It is known since the work of Foster and Vohra (1998) that an $O(T^{2/3})$ rate is achievable for the expected K_1 calibration error (by a randomized algorithm) and it remains the best upper bound for this problem. The best lower bound is $\Omega(T^{0.528})$ shown in the recent work of Qiao and Valiant (2021), a breakthrough improvement over the previous best lower bound of $\Omega(\sqrt{T})$, which can be easily established by generating the true states as unbiased and independent coin flips. Qiao and Zheng (2024) give a non-constructive minimax-based proof that an $O(\sqrt{T})$ upper bound can be achieved for the distance to calibration. They also show an $\Omega(T^{1/3})$ lower bound for the same problem. Soon afterwards, Arunachaleswaran et al. (2024) provide an explicit construction of an efficient algorithm that achieves the $O(\sqrt{T})$ upper bound for the distance to calibration.

Omniprediction. Treating prediction and decision making as separate steps allows us to train a single predictor and use it to solve multiple decision tasks with different utility/loss functions. This separation of training and decision making is the idea behind omniprediction, introduced recently by Gopalan et al. (2022), where the goal is to train a single predictor that allows each downstream decision maker to incur comparable or smaller loss than any alternative decision rule from a benchmark class. Notions from the algorithmic fairness literature (e.g. multicalibration and multiaccuracy (Hebert-Johnson et al., 2018; Kim et al., 2019)) have been used to obtain omnipredictors in various online and offline (batch) settings (Gopalan et al., 2022, 2023a; Hu et al., 2023; Gopalan et al., 2023b, 2024; Garg et al., 2024; Noarov et al., 2023; Kim and Perdomo, 2023). Omniprediction allows better efficiency than training a different model from scratch for each decision task, and it also allows the predictor to be robust to changes in the loss function.

Swap Regret Minimization. Foster and Vohra (1997) first show vanishing swap regret implies convergence to correlated equilibria. In the online learning literature, there is extensive work on swap regret minimization (e.g. Hart and Mas-Colell, 2000, 2001; Blum and Mansour, 2007; Anagnostides et al., 2022, etc). Recently, Peng and Rubinstein (2023); Dagan et al. (2023) prove a lowerbound on the swap regret, which is polynomial in the number of actions. Meanwhile, the calibration literature (Kleinberg et al., 2023; Noarov et al., 2023; Roth and Shi, 2024) differs from the swap regret minimization literature in two aspects: 1) it focuses on developing a robust strategy that minimizes swap regret simultaneously for all decision makers, and 2) it focuses on minimizing swap regret for the special payoff structure of decision tasks. As a special payoff structure, a decision task restricts the adversary to only be able to select a state. Equivalently, the adversary can select payoff from a low rank matrix with the same rank of the state space. Our result implies the lowerbound on swap regret is strictly weaker when there exists a special low-rank constraint on payoff matrix.

Optimization of Scoring Rules. MSR is defined as the maximum swap regret over all decision tasks. Since the payoff in a decision task can be equivalently represented by proper scoring rules (see Section 2.4), the computation of MSR is an optimization problem of scoring rules. A recent literature (Li ⓡ al., 2022; Neyman et al., 2021; Hartline ⓡ al., 2023) studies the optimization

of scoring rules, where Li ⓡ al. (2022) is the most relevant paper. Li ⓡ al. (2022) present two results that are helpful to our problem: 1) under their different optimization objective, the optimal scoring rule can be computed via linear programming, and 2) any bounded scoring rule can be decomposed into a linear combination of V-shaped scoring rules. Following their idea, we design a linear program that computes MSR in polynomial time. Our $O(\sqrt{T}\log T)$ MSR result also uses this linear decomposition of scoring rules (see Section 3.2).

1.3 Paper Organization

We formally introduce the online binary prediction problem, popular measures of the calibration error, and the swap regret of decision tasks in Section 2. We introduce Maximum Swap Regret (MSR) in Section 3 and discuss its alternative formulation using Bregman divergences and its approximation via V-shaped Bregman divergences, which will be useful to establish our main result. Section 4 discusses the connection between MSR and other calibration error metrics. In Section 5, we present our main result, an efficient online binary prediction algorithm that guarantees $O(\sqrt{T}\log T)$ MSR. The key technical idea behind this result is a lemma (Lemma 5.2) we prove in Section 5.1 which allows us to attribute MSR to bucket-wise biases. In Section 6, we discuss the economic utility of calibration from a decision-theoretic perspective, which justifies MSR as a calibration error metric. Additionally, we give a non-constructive but simpler minimax proof of the $O(\sqrt{T}\log T)$ MSR guarantee in Appendix A. This simpler proof also crucially relies on our key technical lemma (Lemma 5.2) in Section 5.1.

2 Preliminaries

Throughout the paper, we use $p \in [0,1]$ to denote a prediction, and use $\theta \in \{0,1\}$ to denote a state. A prediction p can be viewed as a distribution over the state space $\{0,1\}$, and we write $\theta \sim p$ when we sample θ from the Bernoulli distribution with mean p, i.e., $\Pr[\theta = 1] = p$. For a real number x, we use $(x)_+$ or $[x]_+$ to denote $\max\{x, 0\}$. We use $\mathbb{I}[\cdot]$ to denote the 0-1 indicator function: $\mathbb{I}[\text{statement}] = 1$ if the statement is true, and $\mathbb{I}[\text{statement}] = 0$ if the statement is false.

2.1 Online Binary Prediction

We focus on the classic online prediction problem studied by Foster and Vohra (1998). Unless otherwise specified, we present results under the binary prediction setting, i.e. the state space is $\Theta = \{0,1\}$. In this setup, a predictor (algorithm) F makes a prediction $p_t \in [0,1]$ at each round $t = 1, 2, \ldots$, and an adversary A picks a binary state (outcome) $\theta_t \in \{0,1\}$. Both the prediction p_t and the state θ_t can depend on the past history $h_{t-1} = (p_1, \theta_1, \ldots, p_{t-1}, \theta_{t-1})$, but they cannot depend on each other. That is, we can assume without loss of generality that the algorithm chooses p_t and the adversary reveals θ_t simultaneously. The transcript $h_t = (p_1, \theta_1, \ldots, p_t, \theta_t)$ is a function of the strategies of the predictor F and the adversary A. That is, $h_t = h_t(F, A)$.

We allow the predictor to be randomized, in which case we can view its strategy as a distribution $\mathcal{F}$ over deterministic strategies F. When we use an error metric, say MSR, to evaluate the predictions made by our predictor in T rounds, our goal is to minimize the expected value of the error metric w.r.t. the worst-case adversary A, i.e., we want $\mathbf{E}_{F\sim\mathcal{F}}[\mathrm{MSR}(h_T(F, A))]$ to be small for every adversary A.

2.2 Measures of Calibration Error

While there is only one natural and clear definition of perfect calibration, there are various metrics for measuring the calibration error. Foster and Vohra (1998) defines the K_1 calibration error (a.k.a. expected calibration error, ECE), which is a measure of calibration error popularly used in the literature. The K_1 calibration error measures the total absolute distance between the prediction and the empirical distribution. The predictions are restricted to fall in a finite space $Q = \{q_i \in [0,1]\}_i$. Let $n_i = \sum_t \mathbb{I}[p_t = q_i]$ be the count of prediction being q_i in T rounds, and $\widehat{q}_i = \frac{\sum_t \mathbb{I}[p_t = q_i]\theta_t}{n_i}$ be the empirical distribution of the realized state conditioning on prediction is q_i.

Definition 2.1 (K_1 calibration error)**.** *Given a sequence of T predictions $\boldsymbol{p} = (p_t)_{t\in[T]}$ and realizations $\boldsymbol{\theta} = (\theta_t)_{t\in[T]}$ of states. The K_1 calibration error is*

$$K_1(\boldsymbol{p}, \boldsymbol{\theta}) = \sum_{q_i \in Q} n_i |q_i - \widehat{q}_i|.$$

On the upperbound side, Foster and Vohra (1998) gives an algorithm that achieves $O(T^{2/3})$ on K_1 calibration error in worst case of adversary strategy.

Theorem 2.2 ((Foster and Vohra, 1998)(see also Hart, 2022; Foster and Hart, 2021))**.** *There exists a randomized online binary prediction algorithm that guarantees $O(T^{2/3})$ expected K_1 calibration error.*

On the lowerbound side, Qiao and Valiant (2021) show that it is impossible to achieve $\widetilde{O}(T^{1/2})$ K_1 calibration error.

Theorem 2.3 (Qiao and Valiant (2021))**.** *For any randomized online binary prediction algorithm, the expected K_1 calibration error w.r.t. the worst-case adversary is $\Omega(T^{0.528})$.*

The lowerbound above is specific to K_1 calibration error. An alternative metric is the K_2 calibration error, the total squared distance between a prediction and the empirical distribution. Theorem 2.5 shows it is possible to achieve $\widetilde{O}(T^{1/2})$ worst-case expected K_2 error.

Definition 2.4 (K_2 calibration error)**.** *Given a sequence of T predictions $\boldsymbol{p} = (p_t)_{t\in[T]}$ and realizations $\boldsymbol{\theta} = (\theta_t)_{t\in[T]}$ of states. The K_2 calibration error is*

$$K_2(\boldsymbol{p}, \boldsymbol{\theta}) = \sum_{q_i \in Q} n_i (q_i - \widehat{q}_i)^2.$$

Theorem 2.5 (Roth (2022))**.** *There exists a randomized online binary prediction algorithm that guarantees $\widetilde{O}(T^{1/2})$ expected K_2 calibration error.*

In addition to K_1 and K_2 calibration error, we also compare to the smooth calibration error introduced by Kakade and Foster (2008). Unlike K_1 and K_2, the smooth calibration error is continuous in predictions.

Definition 2.6 (Smooth Calibration Error (Kakade and Foster, 2008))**.** *Given a sequence of T predictions $\boldsymbol{p} = (p_t)_{t\in[T]}$ and realizations $\boldsymbol{\theta} = (\theta_t)_{t\in[T]}$ of states. The smooth calibration error is a supremum over the set Σ of 1-Lipschitz functions $\sigma : [0,1] \to [-1,1]$:*

$$\textsc{smCal}(\boldsymbol{p}, \boldsymbol{\theta}) = \sup_{\sigma \in \Sigma} \sum_t \sigma(p_t)(p_t - \theta_t).$$

The definition above is equivalent to K_1 without the 1-Lipschitz constraint on σ. Taking the following non-Lipschitz σ yields K_1.

$$\sigma(q_i) = \begin{cases} 1, & \text{if } q_i - \widehat{q}_i \geq 0; \\ -1, & \text{otherwise.} \end{cases}$$

The smooth calibration error is polynomially related to the distance to calibration, which measures the absolute distance to the closest calibrated prediction.

Definition 2.7 (Distance to Calibration (Błasiok et al., 2023a)). *Given a sequence of T predictions $\boldsymbol{p} = (p_t)_{t\in[T]}$ and realizations $\boldsymbol{\theta} = (\theta_t)_{t\in[T]}$ of states. The distance to calibration is*

$$\textsc{distCal}(\boldsymbol{p}, \boldsymbol{\theta}) = \min_{\widehat{\boldsymbol{p}}: K_1(\widehat{\boldsymbol{p}}, \boldsymbol{\theta})=0} \sum_t |p_t - \widehat{p}_t|.$$

Lemma 2.8 (Błasiok et al. (2023a)). *Smooth calibration error* $\textsc{smCal}$ *is polynomially related to distance to calibration* $\textsc{distCal}$.

$$\frac{\textsc{smCal}}{2T} \leq \frac{\textsc{distCal}}{T} \leq \sqrt{\frac{32\,\textsc{smCal}}{T}}.$$

Qiao and Zheng (2024) prove the existence of an algorithm that achieves $O(\sqrt{T})$ regret on the distance to calibration, following which Arunachaleswaran et al. (2024) give a construction of such an algorithm.

Theorem 2.9 (Qiao and Zheng (2024); Arunachaleswaran et al. (2024)). *There exists a randomized online binary prediction algorithm that guarantees $O(\sqrt{T})$ expected distance to calibration.*

2.3 Sequential Decision Making and Action Swap Regret

Calibrated forecasts imply no swap regret for decision making (Foster and Vohra, 1999). Before introducing predictions and calibration, we define action swap regret in online decision making. In each round t, the agent faces an identical decision task (A, Θ, U). Throughout the paper, we normalize the payoff in the decision task to $[0, 1]$.

- The agent takes action $a \in A$.
- A payoff-relevant state $\theta \in \Theta$ realizes.
- The agent obtains payoff $U : A \times \Theta \to [0, 1]$ as a function of the action and the state.

We define the action swap regret the same as Roth and Shi (2024).

Definition 2.10 (Swap Regret of Actions). *Given a sequence of T actions $\boldsymbol{a} = (a_t)_{t\in[T]}$ and realization $\boldsymbol{\theta} = (\theta_t)_{t\in[T]}$ of states, and fix a decision task with payoff rule U, the action swap regret is*

$$\textsc{ASwap}_U(\boldsymbol{a}, \boldsymbol{\theta}) = \max_{\sigma: A\to A} \sum_{t\in[T]} \Big[U(\sigma(a_t), \theta_t) - U(a_t, \theta_t) \Big].$$

Kleinberg et al. (2023) aim to minimize the external regret for all decision tasks with bounded payoff. We define the external regret here for comparison.

Definition 2.11 (External Regret). *Given a sequence of T actions $\boldsymbol{a} = (a_t)_{t\in[T]}$ and realization $\boldsymbol{\theta} = (\theta_t)_{t\in[T]}$ of states, and fix a decision task with payoff rule U, the external regret is calculated against the best fixed action,*

$$\textsc{Ext}_U(\boldsymbol{a}, \boldsymbol{\theta}) = \max_{a\in A} \Big[U(a, \theta_t) - U(a_t, \theta_t) \Big]. \tag{9}$$

2.4 Proper Scoring Rules and Prediction Swap Regret

Our paper focuses on the prediction swap regret, a stronger swap regret than the action swap regret. When the agent is assisted with a sequence of predictions, the agent obtains payoff by acting in response to the prediction. If the prediction is the true distribution where the state is drawn, the best response a^* is a function of the prediction:

$$a^*(p) = \arg\max_{a \in A} \mathbf{E}_{\theta \sim p} \left[U(a, \theta) \right]. \tag{10}$$

When the agent best responds to a prediction, they have the prediction swap regret. The modification rule σ allows the agent to modify the action as a response to the prediction.

Definition 2.12 (Prediction Swap Regret)**.** *Given a sequence of T predictions $\boldsymbol{p} = (p_t)_{t \in [T]}$ and realization $\boldsymbol{\theta} = (\theta_t)_{t \in [T]}$ of states, and fix a decision task with payoff rule U, the prediction swap regret is*

$$\text{PSwap}_U(\boldsymbol{p}, \boldsymbol{\theta}) = \max_{\sigma : \Delta(\Theta) \to A} \sum_{t \in [T]} \Big[U(\sigma(p_t), \theta_t) - U(a^*(p_t), \theta_t) \Big],$$

where a^* is the best response to prediction as defined in Equation (10).

For ease of notation, we present the prediction swap regret in a scoring rule S, representing the utility as a function of the prediction instead of the action.

Definition 2.13 (Scoring rule induced by decision task)**.** *Given a decision task (A, Θ, U) and its corresponding best response function a^*, we define an induced scoring rule $S_U : \Delta(\Theta) \times \Theta \to \mathbb{R}$ such that for any prediction $p \in \Delta(\Theta)$ and state $\theta \in \Theta$,*

$$S_U(p, \theta) = U(a^*(p), \theta).$$

That is, $S_U(p, \theta)$ is the payoff of a decision maker who chooses the best response $a^*(p)$ based on the prediction p.

Such scoring rules induced by a decision task are equivalent to the class of proper scoring rules (Frongillo and Kash, 2014). Thus, the alternative definition of prediction swap regret considers swapping predictions for proper scoring rules (Claim 2.16).

Proper scoring rules are defined such that the expected score is maximized when the prediction is the true distribution. Evaluated with a proper scoring rule, the predictor has an incentive to truthfully report the distribution in order to maximize expected score.

Definition 2.14 (Proper Scoring Rule)**.** *A scoring rule S is proper if for any $p' \in \Delta(\Theta)$ that is not the true distribution θ is generated,*

$$\mathbf{E}_{\theta \sim p} \left[S(p, \theta) \right] \geq \mathbf{E}_{\theta \sim p} \left[S(p', \theta) \right].$$

Claim 2.15 shows that 1) the scoring rule induced by a decision task is a proper scoring rule, and 2) any proper scoring rule measures the best-response payoff in a decision task.

Claim 2.15. *Proper scoring rules and payoffs in decision task are equivalent:*

1. *Any scoring rule S_U induced by a decision task (A, Θ, U) is proper.*
2. *For any proper scoring rule S, there exists a decision task with payoff U, such that $U(a^*(p), \theta) = S(p, \theta)$ for every prediction p.*

Proof. 2 is straight forward by defining $A = \Delta(\Theta)$, the action space as the prediction space.

To see 1, notice that the score is the best-response score with $S_U(p, \theta) = U(a^*(p), \theta)$. Hence, if the prediction is $p' \neq p$ not the distribution where θ is drawn, the agent obtains a payoff weakly lower than the payoff by best responding to the true distribution.

$$\mathbf{E}_{\theta\sim p}\left[S(p', \theta)\right] = \mathbf{E}_{\theta\sim p}\left[U(a^*(p'), \theta)\right] \leq \mathbf{E}_{\theta\sim p}\left[U(a^*(p), \theta)\right] = \mathbf{E}_{\theta\sim p}\left[S(p, \theta)\right]. \qquad \square$$

The prediction swap regret is equivalently the regret from swapping predictions, which was first introduced in Foster and Vohra (1998) with the quadratic scoring rule $S(p, \theta) = 1 - (p - \theta)^2$.

Claim 2.16. *Given a sequence of T predictions $\boldsymbol{p} = (p_t)_{t\in[T]}$ and realizations $\boldsymbol{\theta} = (\theta_t)_{t\in[T]}$ of states, and fix a decision task with induced proper scoring rule S. The prediction swap regret PSWAP_U is equivalently*

$$\text{SWAP}_S(\boldsymbol{p}, \boldsymbol{\theta}) = \max_{\sigma:\Delta(\Theta)\to\Delta(\Theta)} \sum_{t\in[T]} \Big[S(\sigma(p_t), \theta_t) - S(p_t, \theta_t)\Big].$$

The prediction swap regret is stronger than the action swap regret. If an algorithm generates a sequence of predictions with low prediction swap regret, then the agent has low action swap regret if they best respond to the predictions. To see this, notice that the modification rule in prediction swap regret has more power. The prediction swap regret allows the agent to modify the action conditioning on each prediction. If two predictions have the same best-response action, the action swap regret does not allow the agent to apply different modification to the two predictions.

Claim 2.17. *Given a decision task with payoff rule U, the corresponding scoring rule is denoted S. If the agent best responds by taking $a_t = a^*(p_t)$ at each round t,*

$$\text{ASWAP}_U(\boldsymbol{a}, \boldsymbol{\theta}) \leq \text{PSWAP}_U = \text{SWAP}_S(\boldsymbol{p}, \boldsymbol{\theta}).$$

Claim 2.18 ((Kleinberg et al., 2023, Theorem 12))**.** [3] *For any proper scoring rule $S : [0, 1]\times\{0, 1\} \to [0, 1]$ and any sequences $\boldsymbol{p} = (p_1, \ldots, p_T) \in [0, 1]^T, \boldsymbol{\theta} = (\theta_1, \ldots, \theta_T) \in \{0, 1\}^T$, it holds that*

$$\text{SWAP}_S(\boldsymbol{p}, \boldsymbol{\theta}) \leq 2K_1(\boldsymbol{p}, \boldsymbol{\theta}).$$

In fact, K_2 can be written as the prediction swap regret for quadratic scoring rule (a.k.a. squared loss), while K_1 cannot be represented by the swap regret on any proper scoring rule.

Lemma 2.19. *Define the quadratic scoring rule $S_2(p, \theta) = 1 - (p - \theta)^2$. We have*

$$K_2 = \text{SWAP}_{S_2}.$$

Lemma 2.20. *There does not exist a proper scoring rule S, such that for any sequence of predictions $\boldsymbol{p}$ and states $\boldsymbol{\theta}$,*

$$K_1(\boldsymbol{p}, \boldsymbol{\theta}) = \text{SWAP}_S(\boldsymbol{p}, \boldsymbol{\theta}).$$

Lemma 2.19 follows immediately from the definitions of K_2 and S_2, whereas Lemma 2.20 can be easily proved using the Bregman divergence characterization of proper scoring rules (see Proposition 3.6).

Kleinberg et al. (2023) defines the U-calibration error, the maximum external regret when the decision maker best responds to the predictions.

[3] Kleinberg et al. (2023) assume that the output of the scoring rule S is in $[-1, 1]$, whereas we assume the output is in $[0, 1]$. Thus, the constant 4 in their bound translates to the constant 2 here.

Definition 2.21 (U-calibration Error)**.** *Let $\boldsymbol{p} = (p_t)_{t\in[T]}$ be a sequence of T predictions and let $\boldsymbol{\theta} = (\theta_t)_{t\in[T]}$ be the realization of states. Consider a decision maker with payoff function $U : A\times\{0,1\} \to [0,1]$ who best responds to the predictions by taking $a_t = a^*(p_t) = \arg\max_{a\in A} \mathbf{E}_{\theta\sim p_t}[U(a,\theta)]$. Let $\boldsymbol{a}_U = (a_1,\dots,a_T)$ be the vector of best-response actions. The U-calibration error is the maximum external regret over decision tasks with bounded payoff in $[0,1]$:*

$$\textsc{UCal}(\boldsymbol{p},\boldsymbol{\theta}) = \sup_U \textsc{Ext}_U(\boldsymbol{a}_U,\boldsymbol{\theta}),$$

where the supremum is over all payoff functions $U : A\times\{0,1\}\to[0,1]$ with arbitrary action spaces A.

3 Maximum Swap Regret

In this section, we introduce our central notion: Maximum Swap Regret (MSR). In Sections 3.1 and 3.2, we discuss connections between MSR and Bregman divergences that will be useful for obtaining our main result in Section 5. We show an efficient algorithm for computing MSR in Section 3.3.

Definition 3.1 (Maximum Swap Regret)**.** *Given a sequence of T predictions $\boldsymbol{p} = (p_t)_{t\in[T]}$ and realization $\boldsymbol{\theta} = (\theta_t)_{t\in[T]}$ of states, we define the Maximum Swap Regret (MSR) as*

$$\mathrm{MSR}(\boldsymbol{p},\boldsymbol{\theta}) = \sup_{S\in[0,1]} \textsc{Swap}_S(\boldsymbol{p},\boldsymbol{\theta}),$$

where the supremum is over all proper scoring rules S with range bounded in $[0,1]$.

The MSR as defined above is equal to the maximum action swap regret (Definition 2.10) of best responding decision makers with payoffs bounded in $[0,1]$:

Lemma 3.2. *Let $\boldsymbol{p} = (p_t)_{t\in[T]}$ be a sequence of T predictions and let $\boldsymbol{\theta} = (\theta_t)_{t\in[T]}$ be the realization of states. Consider a decision maker with payoff function $U : A\times\{0,1\}\to[0,1]$ who best responds to the predictions by taking $a_t = a^*(p_t) = \arg\max_{a\in A}\mathbf{E}_{\theta\sim p_t}[U(a,\theta)]$. Let $\boldsymbol{a}_U = (a_1,\dots,a_T)$ be the vector of best-response actions. We have*

$$\mathrm{MSR}(\boldsymbol{p},\boldsymbol{\theta}) = \sup_U \textsc{ASwap}_U(\boldsymbol{a}_U,\boldsymbol{\theta}),$$

where the supremum is over all payoff functions $U : A\times\{0,1\}\to[0,1]$ with arbitrary action spaces A.

Proof. For any decision task with payoff function U, the prediction swap regret for the corresponding scoring rule S is higher than the action swap regret: $\textsc{ASwap}_U(\boldsymbol{a}_U,\boldsymbol{\theta}) \le \textsc{Swap}_S(\boldsymbol{p},\boldsymbol{\theta})$ (Claim 2.17). Thus, $\mathrm{MSR} \ge \sup_U \textsc{ASwap}_U(\boldsymbol{a}_U,\boldsymbol{\theta})$. On the other hand, for any scoring rule S, construct a decision task with payoff U as in the proof of Claim 2.15, by setting $A = [0,1]$. The resulting decision task has $\textsc{Swap}_S = \textsc{ASwap}_U$. Thus we have $\mathrm{MSR} = \sup_U \textsc{ASwap}_U(\boldsymbol{a}_U,\boldsymbol{\theta})$. □

3.1 Bregman Divergence Formulation

As an important preparation for our main result in Section 5, we show that the prediction swap regret can be written in an equivalent form using the Bregman divergence (Proposition 3.6). This follows from the convex function formulation of proper scoring rules from McCarthy (1956); Savage (1971).

First, the swap regret is maximized when predictions are swapped to the conditional empirical distributions.

Lemma 3.3. *Given a sequence of T predictions $\boldsymbol{p} = (p_t)_{t\in[T]}$ from a finite set $Q = \{q_1, \dots, q_m\} \subseteq [0,1]$ and realizations $\boldsymbol{\theta} = (\theta_t)_{t\in[T]}$ of states, define n_i and $\widehat{q}_i$ as in (2) and (3). For any proper scoring rule S,*

$$\textsc{Swap}_S(\boldsymbol{p}, \boldsymbol{\theta}) = \sum_{i\in[m]} n_i \sum_{j\in[T]} \Big(S(\widehat{q}_i, \theta_t) - S(q_i, \theta_t) \Big) \mathbb{I}\left[p_t = q_i\right],$$

where $n_i = \sum_t \mathbb{I}\left[p_t = q_i\right]$ is the count of prediction being q_i in T rounds, and $\widehat{q}_i = \frac{\sum_t \mathbb{I}[p_t = q_i]\theta_t}{n_i}$ is the empirical distribution of the realized state conditioning on prediction is q_i.

Lemma 3.3 follows directly from the properness of scoring rules. Conditioning on the prediction is q_i, the state follows the empirical distribution:

$$\frac{1}{n_i} \sum_{t\in[T]} S(p, \theta_t) \mathbb{I}\left[p_t = q_i\right] = \mathbf{E}_{\theta\sim\widehat{q}_i}\left[S(p,\theta)\right].$$

By properness of the scoring rule, predicting the empirical distribution maximizes the expected score (average score).

Fixing a prediction q_i, we show there exists a Bregman divergence formulation of $\sum_{t\in[T]}(S(\widehat{q}_i, \theta_t) - S(q_i, \theta_t))\mathbb{I}\left[p_t = q_i\right]$. Each proper scoring rule induces a Bregman divergence, measuring the loss from incorrect prediction. The Bregman divergence is defined with the convex utility function of a scoring rule. See Figure 1 for a geometric demonstration of proper scoring rules.

Theorem 3.4 (McCarthy (1956); Savage (1971))**.** *A scoring rule is proper if and only if there exists a convex utility function $u : \Delta(\Theta) \to \mathbb{R}$ such that*

$$S(p, \theta) = u(p) + \nabla u(p) \cdot (\theta - p). \tag{11}$$

Notice that when the prediction is the true distribution, $\mathbf{E}_{\theta\sim p}\left[S(p,\theta)\right] = u(p)$ since $\mathbf{E}\left[\theta\right] = p$.

Definition 3.5 (Bregman divergence)**.** *Specified by a convex function $u : \Delta(\Theta) \to \mathbb{R}$, the Bregman divergence $\textsc{Breg}(p', p)$ is defined as*[4]

$$\textsc{Breg}(p, \widehat{p}) = u(\widehat{p}) - u(p) + \nabla u(p)(p - \widehat{p}). \tag{12}$$

Given a proper scoring rule S, we write the Bregman divergence defined with $u(p) = \mathbf{E}_{\theta\sim p}\left[S(p,\theta)\right]$ as $\textsc{Breg}_S$. It follows that

$$\frac{1}{n_i} \sum_{t\in[T]} (S(\widehat{q}_i, \theta_t) - S(q_i, \theta_t))\mathbb{I}\left[p_t = q_i\right] = \mathbf{E}_{\theta\sim\widehat{q}_i}\left[S(\widehat{q}_i, \theta) - S(q_i, \theta)\right] = \textsc{Breg}_S(q_i, \widehat{q}_i).$$

Proposition 3.6. *Given a sequence of T predictions $\boldsymbol{p} = (p_t)_{t\in[T]}$ from a finite set $Q = \{q_1, \dots, q_m\} \subseteq [0,1]$ and realizations $\boldsymbol{\theta} = (\theta_t)_{t\in[T]}$ of states, define n_i and $\widehat{q}_i$ as in (2) and (3). For any proper scoring rule S,*

$$\textsc{Swap}_S(\boldsymbol{p}, \boldsymbol{\theta}) = \sum_{i\in[m]} n_i \textsc{Breg}_S(q_i, \widehat{q}_i). \tag{13}$$

[4]The definition of Bregman divergence here reverses the order of input in the conventional definition to align with proper scoring rules. In the conventional definition, $\textsc{Breg}(p, \widehat{p}) = u(p) - u(\widehat{p}) + \nabla u(\widehat{p})(\widehat{p} - p)$.

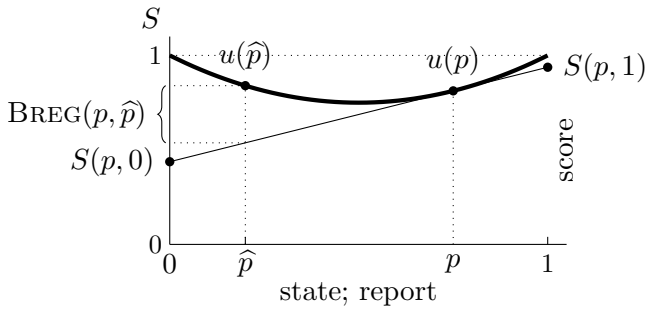

Figure 1: The graphic explanation of the connection between proper scoring rule and Bregman divergence. The thick convex curve plots the convex utility function $u(p)$ for a proper scoring rule. Fix a report, the score $S(p,\theta) = u(p) + \nabla u(p)(\theta - p)$ is the extreme points on the gradient hyperplane passing $u(p)$ (the thin line). Given empirical distribution $\widehat{p}$, the Bregman divergence $\textsc{Breg}(p,\widehat{p})$ is the loss of reporting p instead of $\widehat{p}$.

3.2 Approximation via V-Bregman Divergences

Following ideas from Li ⓡ al. (2022); Kleinberg et al. (2023), we decompose any Bregman divergence into a linear combination on a basis. Thus, our MSR is bounded by the maximum swap regret on the basis, the V-Bregman divergences.

Definition 3.7 (V-Bregman divergence)**.** *The V-Bregman divergence with kink μ is defined as*

$$\textsc{VBreg}_\mu(q,\widehat{q}) = \begin{cases} 0, & \text{if 1) } q < \mu \text{ and } \widehat{q} \le \mu \text{ or 2)} q \ge \mu \text{ and } \widehat{q} \ge \mu; \\ \frac{|\widehat{q}-\mu|}{\max\{1-\mu,\mu\}}, & \text{otherwise.} \end{cases} \tag{14}$$

We note that any V-Bregman divergence can be induced by a V-shaped proper scoring rule bounded in $[0,1]$. More specifically, the V-shaped scoring rule offers two actions for the agent to decide, with optimal decision rule as a threshold at μ. Without loss of generality, suppose $\mu \le 1/2$. $\textsc{VBreg}_\mu$ can be induced from the following proper scoring rule, which is shown in Figure 2.

$$S_\mu(p,\theta) = \begin{cases} 3/4 - \frac{1}{2}\cdot\frac{\theta-\mu}{1-\mu} & \text{if } p \le \mu \\ 1/4 + \frac{1}{2}\cdot\frac{\theta-\mu}{1-\mu} & \text{else,} \end{cases} \tag{15}$$

The following lemma allows us to decompose a general Bregman divergence as a linear combination of V-Bregman divergences.

Lemma 3.8 (Li ⓡ al. (2022); Kleinberg et al. (2023))**.** *Let $u : [0,1] \to [0,1]$ be a twice differentiable convex function. Let S be the proper scoring rule generated by u as in (11). Then for every $p, \widehat{p} \in [0,1]$,*

$$\textsc{Breg}_S(p,\widehat{p}) = \int_{\mu=0}^{1} u''(\mu)\cdot\max(1-\mu,\mu)\textsc{VBreg}_\mu d\mu. \tag{16}$$

Definition 3.9 (V-swap Regret)**.** *Given a sequence of T predictions $\boldsymbol{p} = (p_t)_{t\in[T]}$ from a finite set $Q = \{q_1,\dots,q_m\} \subseteq [0,1]$ and realizations $\boldsymbol{\theta} = (\theta_t)_{t\in[T]}$ of states, the V-swap regret is defined as the supremem regret over all V-Bregman divergences:*

$$\textsc{VSwap} = \sup_{\mu\in[0,1]} \sum_{i\in[m]} n_i \textsc{VBreg}_\mu(q_i,\widehat{q}_i). \tag{17}$$

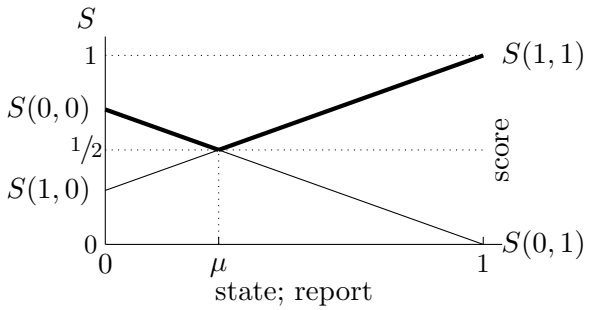

Figure 2: The thick black line plots the special convex utility function u for the scoring rule. The convex utility function is V-shaped, consisting of two linear pieces intersecting at μ. Once fixing the prediction p, the score $S(p, \theta) = u(p) + \nabla u(p) \cdot (\theta - p)$ is linear in the state θ. The scoring rule offers two set of scores for the prediction to selecttwo linear lines for prediction. When $p \le q_0$, the prediction selects scores $\{S(0,0), S(0,1)\}$. Otherwise, the prediction selects $\{S(1,0), S(1,1)\}$.

Theorem 3.10. *V-swap regret is a constant-factor approximation of MSR. That is, for any sequence of predictions $\boldsymbol{p} = (p_1, \dots, p_T) \in [0,1]^T$ and any sequence of states $\boldsymbol{\theta} = (\theta_1, \dots, \theta_T) \in \{0,1\}^T$,*

$$\textsc{VSwap}(\boldsymbol{p}, \boldsymbol{\theta}) \le \textsc{MSR}(\boldsymbol{p}, \boldsymbol{\theta}) \le 2\textsc{VSwap}(\boldsymbol{p}, \boldsymbol{\theta}).$$

Proof. Since $\textsc{VBreg}_\mu$ is a special case of a Bregman divergence $\textsc{Breg}_S$ for a scoring rule S, comparing the definition of VSwap in (17) and the expression of Swap in (13), we immediately get $\textsc{VSwap}(\boldsymbol{p}, \boldsymbol{\theta}) \le \textsc{MSR}(\boldsymbol{p}, \boldsymbol{\theta})$.

Now let us consider any proper scoring rule $S : [0,1] \times \{0,1\} \to [0,1]$. By Theorem 3.4, there exists a convex function $u : [0,1] \to [0,1]$ with bounded sub-gradients $\nabla u(p) \in [-1, 1]$ such that (11) holds. We can construct a sequence of twice-differentiable convex functions $u_1, u_2, \dots$ such that for every $p \in [0,1]$, $u_j(p) \to u(p)$ and $u_j'(p) \to \nabla u(p)$ as $j \to +\infty$. Let S_j denote the proper scoring rule generated by u_j as in (11). For every $p, \widehat{p} \in [0,1]$, we have

$$\lim_{j \to \infty} \textsc{Breg}_{S_j}(p, \widehat{p}) = \textsc{Breg}_S(p, \widehat{p}).$$

Let $Q = \{q_1, \dots, q_m\} \subseteq [0,1]$ be a finite prediction space such that $p_t \in Q$ for every $t \in [T]$. Define n_i and $\widehat{q}_i$ as in (2) and (3). By Lemma 3.8,

$$\begin{aligned}
\sum_{i \in [m]} n_i \textsc{Breg}_{S_j}(q_i, \widehat{q}_i) &= \sum_{i \in [m]} n_i \int_{\mu=0}^{1} u_j''(\mu) \cdot \max(1-\mu, \mu) \textsc{VBreg}_\mu(q_i, \widehat{q}_i) \mathrm{d}\mu \\
&= \int_{\mu=0}^{1} u_j''(\mu) \cdot \max(1-\mu, \mu) \sum_{i \in [m]} n_i \textsc{VBreg}_\mu(q_i, \widehat{q}_i) \mathrm{d}\mu \\
&\le \int_{\mu=0}^{1} u_j''(\mu) \cdot \max(1-\mu, \mu) \textsc{VSwap}(\boldsymbol{p}, \boldsymbol{\theta}) \mathrm{d}\mu \\
&= \textsc{VSwap}(\boldsymbol{p}, \boldsymbol{\theta}) \int_{\mu=0}^{1} u_j''(\mu) \cdot \max(1-\mu, \mu) \mathrm{d}\mu \\
&\le \textsc{VSwap}(\boldsymbol{p}, \boldsymbol{\theta}) \int_{\mu=0}^{1} u_j''(\mu) \cdot \mathrm{d}\mu \\
&= \textsc{VSwap}(\boldsymbol{p}, \boldsymbol{\theta})(u_j'(1) - u_j'(0)).
\end{aligned}$$

The right hand side converges as $j \to \infty$:

$$\lim_{j\to\infty} \textsc{VSwap}(\boldsymbol{p},\boldsymbol{\theta})(u_j'(1)-u_j'(0)) = \textsc{VSwap}(\boldsymbol{p},\boldsymbol{\theta})(\nabla u(1)-\nabla u(0)) \le 2\textsc{VSwap}(\boldsymbol{p},\boldsymbol{\theta}).$$

Therefore, by Proposition 3.6,

$$\begin{aligned}\textsc{Swap}_S(\boldsymbol{p},\boldsymbol{\theta}) = \sum_{q_i\in Q} n_i \textsc{Breg}_S(q_i,\widehat{q}_i) &= \lim_{j\to\infty}\sum_{q_i\in Q} n_i \textsc{Breg}_{S_j}(q_i,\widehat{q}_i)\\ &\le \lim_{j\to\infty} \textsc{VSwap}(\boldsymbol{p},\boldsymbol{\theta})(u_j'(1)-u_j'(0))\\ &\le 2\textsc{VSwap}(\boldsymbol{p},\boldsymbol{\theta}).\end{aligned}$$

Taking supremum over S proves $\mathrm{MSR}(\boldsymbol{p},\boldsymbol{\theta}) \le 2\textsc{VSwap}(\boldsymbol{p},\boldsymbol{\theta})$. □

3.3 Computation of Maximum Swap Regret

We allow the state space to be non-binary in this section. By solving a linear program, the MSR can be computed in time polynomial in the size of the prediction space $|Q|$ and the state space $|\Theta|$.

Theorem 3.11. *Given a sequence of predictions $\boldsymbol{p}$ and states $\boldsymbol{\theta}$, suppose Q is the space of predictions.* MSR *can be computed in time polynomial in $|Q|$ and $|\Theta|$.*

Proof. The computation of MSR is an optimization problem over the space of proper scoring rules. We follow the idea in Li ⓡ al. (2022); Kleinberg et al. (2023). The optimal scoring rule can be computed by solving a linear program. Let $\widehat{Q} = \{\widehat{q}_i\}_i$ be the set of empirical distributions for each $q_i \in Q$. Define $\widehat{\Theta} = \{\widetilde{\theta} \mid j \in |\Theta|, \widetilde{\theta}_j = 1, \widetilde{\theta}_{j'\neq j} = 0\}$ as the set of indicator predictions of a certain state. Define the space of predictions as $\mathcal{Q} = Q \cup \widehat{Q} \cup \widehat{\Theta}$. We set the scores $s_{q,\theta}, \forall q \in \mathcal{Q}, \theta \in \Theta$ as variables in the linear program.

$$\begin{aligned}\max_{S} \quad & \sum_{q\in Q,\theta\in\Theta} (s_{\widehat{q},\theta} - s_{q,\theta})\widehat{q}(\theta) \\ \text{s.t.} \quad & s_{q,\theta} \in [0,1], \quad \forall q \in \mathcal{Q}, \theta \in \Theta && \text{(bounded payoff)}\\ & \sum_{\theta} q(\theta) s_{q,\theta} \ge \sum_{\theta} q(\theta) s_{q',\theta}, \quad \forall q, q' \in \mathcal{Q} && \text{(properness)}\end{aligned}$$

The following proper scoring rule achieves the mamximum swap regret.

$$S(p,\theta) = s_{q,\theta}, \text{ where } q = \arg\max_{q'\in\mathcal{Q}} \mathbf{E}_{\theta\sim p}\left[s_{q',\theta}\right].$$

□

4 Maximum Swap Regret and Calibration Errors

In this section, we discuss connections of our MSR to calibration errors in the literature. We show that both K_1 and K_2 are polynomially related to MSR (Theorem 4.1), but neither is a constant-factor approximation (Example 4.2). We give examples where the smooth calibration error differs significantly from MSR in either direction (Example 4.3). We also show the U-calibration error lowerbounds MSR (Proposition 4.4), but is not polynomially related (Example 4.5).

Theorem 4.1. *For any sequence of predictions and states,*

$$\begin{aligned}\left(\frac{K_1}{T}\right)^2 \le \frac{\mathrm{MSR}}{T} &\le \frac{2K_1}{T},\\ \frac{K_2}{T} \le \frac{\mathrm{MSR}}{T} &\le 2\sqrt{\frac{K_2}{T}}.\end{aligned}$$

Proof of Theorem 4.1. On the upper bound side, by Claim 2.18, any proper scoring rule S with range bounded in $[0,1]$ satisfies $\textsc{Swap}_S \le 2K_1$, implying $\frac{\mathrm{MSR}}{T} \le \frac{2K_1}{T}$. It is easy to check that $\frac{K_1}{T} \le \sqrt{\frac{K_2}{T}}$ (see e.g. Kleinberg et al., 2023; Roth, 2022). Thus we get the other upper bound $\frac{\mathrm{MSR}}{T} \le 2\sqrt{\frac{K_2}{T}}$.

On the lower bound side, by Lemma 2.19, we have $\frac{\mathrm{MSR}}{T} \ge \frac{K_2}{T}$. Combining this with the fact $\frac{K_1}{T} \le \sqrt{\frac{K_2}{T}}$ yields $\frac{\mathrm{MSR}}{T} \ge (\frac{K_1}{T})^2$. □

All four inequalities in Theorem 4.1 are tight up to constant factors. We demonstrate the tight examples in Example 4.2.

Example 4.2. *Consider the following two miscalibrated predictors.*

(a) (Tight example for upper bounds of MSR*) The state is deterministically* 1*. The predictor deterministically predicts* $1-\epsilon$*.*

In this case, $K_1 = \epsilon T$*,* $K_2 = \epsilon^2 T$*,* $\mathrm{MSR} = \Theta(\epsilon T)$*.*

(b) (Tight example for lower bounds of MSR*) The* T *rounds are divided into* $\sqrt{T}$ *periods, each with* $\sqrt{T}$ *rounds. In each period* i*, the empirical distribution of the state is* $\frac{i}{\sqrt{T}}$*, and the predictor predicts* $\frac{i}{\sqrt{T}} + \frac{1}{\sqrt{T}}$*.*

In this case, $K_1 = \sqrt{T}$*,* $K_2 = 1$*,* $\mathrm{MSR} \in [1, 8]$*.*

From Example 4.2, we see that when predictions concentrate in a small interval, K_1 calculates the swap regret in the correct order as in (a). However, if predictions have high variance as in (b), the calibration error does not simply add up to the total loss in decision. Consider the V-Bregman divergence which corresponds to a decision problem with two actions. Suppose the kink, also the decision threshold, is at $1/2$. Miscalibration at extreme predictions near 0 or 1 will not induce a swap regret to the agent. This intuition is explained by Lemma 5.3, with which we prove the $\widetilde{O}(\sqrt{T})$ expected regret later. We state it here and prove it in Section 5.1.

Define the bias in bucket i (i.e. conditional K_1 calibration error on prediction q_i):

$$G_i = n_i |q_i - \widehat{q}_i|.$$

Lemma 5.3. *Let* T, m *be positive integers. Define* $Q = \{q_1, \dots, q_m\} \subseteq [0,1]$ *where* $q_i = i/m$ *for every* $i = 1, \dots, m$*. Given a sequence of predictions* $\boldsymbol{p} = (p_1, \dots, p_T) \in Q^T$ *and realized states* $\boldsymbol{\theta} = (\theta_1, \dots, \theta_T) \in \{0,1\}^T$*, define* n_i *and* $\widehat{q}_i$ *as in (2) and (3). Define* $G_i := n_i|\widehat{q}_i - q_i|$*. Fix a V-Bregman divergence with kink* μ*, the prediction swap regret is bounded by*

$$\textsc{Swap}_\mu(\boldsymbol{p}, \boldsymbol{\theta}) \le 2 \sum_{i \in [m]} \left(G_i - n_i |q_i - \mu|\right)_+ .$$

Proof of Example 4.2. The calculation of K_1 and K_2 are straightforward. We show the calculation of MSR separately for (a) and (b).

(a) By Lemma 5.3, $\mathrm{MSR} \le 2\mathrm{VSwap} \le 4\epsilon T$. We can find a Bregman divergence such that $\textsc{Swap} \ge \frac{\epsilon T}{2}$. Consider the scoring rule S that has a V-Bregman divergence with kink $\mu = 1 - \frac{\epsilon}{2}$.

$$\textsc{Swap} = T \frac{|1-\mu|}{\mu} \ge \frac{\epsilon T}{2}.$$

Thus, $\mathrm{MSR} = \Theta(\epsilon T)$.

(b) We can prove VSwap $= \Theta(1)$. Fix any V-Bregman divergence with kink μ, consider the corresponding swap regret Swap_μ.

$$\text{Swap}_\mu = \sqrt{T}\sum_{i=1}^{\sqrt{T}} \frac{|\frac{i}{\sqrt{T}} - \mu|}{\max\{\mu, 1-\mu\}} \left(\mathbb{I}\left[\frac{i}{\sqrt{T}} > \mu > \frac{i+1}{\sqrt{T}}\right] + \mathbb{I}\left[\frac{i}{\sqrt{T}} < \mu < \frac{i+1}{\sqrt{T}}\right]\right)$$

We notice that for predictions that induces a non-zero swap regret, it must be $|\frac{i}{\sqrt{T}} - \mu| \le \frac{1}{\sqrt{T}}$. Since $\max\{\mu, 1-\mu\} \ge \frac{1}{2}$,

$$\text{Swap}_\mu \le \sqrt{T} \cdot 2 \cdot \frac{1}{1/2} \frac{1}{\sqrt{T}} = 4.$$

We know MSR $\le 2\text{VSwap} \le 8$.

□

MSR is not polynomially related to smCal or distCal. Specifically, example (b) in Example 4.3 shows the regret bound of $\widetilde{O}(\sqrt{T})$ of distCal does not apply to MSR.

Example 4.3. *We give two examples of miscalibrated predictors.*

(a) (Large distCal*, small* MSR*) The same example as (b) in Example 4.2.*

distCal $\ge$ smCal $\ge \sqrt{T}$, MSR $\in [1, 2]$.

(b) (Small distCal*, large* MSR*) At the first $\frac{T}{2}$ rounds, $\theta = 1$ deterministically, and the predictor predicts $1/2 + \epsilon$. At the later $\frac{T}{2}$ rounds, $\theta = 0$ deterministically, and the predictor predicts $1/2 - \epsilon$.*

smCal $\le$ distCal $\le \epsilon T$, MSR $= \Omega(T)$.

We can take $\epsilon = \frac{1}{\sqrt{T}}$, which can be arbitrarily small.

Proof of Example 4.3. We calculate distCal and MSR separately for each example.

(a) It only remains to show smCal $\ge \sqrt{T}$. By Definition 2.6, take Lipschitz function $\sigma(\cdot) = 1$.

$$\text{smCal} \ge \sum_t (p_t - \theta_t) = T \cdot \frac{1}{\sqrt{T}} = \sqrt{T}.$$

(b) For distCal, a calibrated predictor always predicts $1/2$.

$$\text{distCal} \le \sum_{t\in[T]} |p_t - \frac{1}{2}| = T\epsilon.$$

For MSR, consider the V-Bregman divergence with kink $1/2 + 2\epsilon$. The swap regret for this V-Bregman divergence is

$$\text{Swap}_{1/2} \ge \frac{T}{2}(1 - 1/2 - 2\epsilon) = (1/4 - \epsilon)T.$$

MSR $\ge \text{Swap}_{1/2}$ implies MSR $= \Omega(T)$.

□

Vanishing U-calibration error is necessary but not sufficient for calibration. By definition, the U-calibration error lowerbounds MSR.

Proposition 4.4. *For any sequence of predictions and states,*

$$\text{UCAL} \leq \text{MSR}.$$

Kleinberg et al. (2023) gives an example where the U-calibration error is 0, while K_1, K_2 and smCAL are non-zero. We present the same example here and show MSR is linear in T.

Example 4.5. *The following example has* 0 *U-calibration error but* MSR *linear in* T.

At the first $\frac{T}{2}$ *rounds,* $\theta = 1$ *deterministically; at the later* $\frac{T}{2}$ *rounds,* $\theta = 0$ *deterministically. The predictor predicts the empirical frequency of the state. I.e. the predictor predicts* $p_t = 1$ *for* $t \leq \frac{T}{2}$*, and* $p_t = \frac{T}{2t}$ *for* $t > \frac{T}{2}$.

In this example, $\lim_{T\to\infty} \text{UCAL} = 0$, $\text{MSR} = \Omega(T)$.

Proof of Example 4.5. Kleinberg et al. (2023) have shown $\lim_{T\to\infty} \text{UCAL} = 0$. We now prove $\text{MSR} = \Omega(T)$ by calculating the swap regret for the V-Bregman divergence. The difference between the external regret and swap regret for the V-Bregman divergence and the corresponding V-shaped scoring rule illustrates the reason why vanishing UCAL does not imply calibration. Consider a V-Bregman divergence with kink $\mu \geq 1/2$ and its corresponding V-shaped scoring rule S_μ, similar as Equation (15):

$$S_\mu(p, \theta) = \begin{cases} 3/4 - \frac{1}{2} \cdot \frac{\mu - \theta}{\mu} & \text{if } p \geq \mu \\ 1/4 + \frac{1}{2} \cdot \frac{\mu - \theta}{\mu} & \text{else,} \end{cases} \tag{18}$$

First, consider the external regret for the decision task with action space $A = [0, 1]$ and payoff $U(a, \theta) = S_\mu(p, \theta)$. The best fixed action is the empirical distribution $1/2$, achieving a score (payoff)

$$\sum_{t\in[T]} S_\mu(1/2, \theta_t) = \sum_{t\in[T]} \left(1/4 + \frac{1}{2} \cdot \frac{\mu - \theta_t}{\mu}\right) = \frac{3T}{4} - \frac{T}{4\mu}.$$

We are going to show, the predictor in the example achieves a payoff higher than always predicting $1/2$.

$$\begin{aligned} \sum_{t\in[T]} S_\mu(p_t, \theta_t) &= \sum_{t\leq T/2\mu} S_\mu(p_t, 1) + \sum_{t>T/2\mu}^{T} S_\mu(\frac{T}{2t}, 0) \\ &= \frac{T}{2\mu}(1/4 + 1/2\mu) + 3/4(T - \frac{T}{2\mu}) \\ &= \frac{3T}{4} - \frac{T}{4\mu} + \frac{T}{4\mu^2}. \end{aligned}$$

The external regret for S_μ is non-positive.

However, the swap regret is high. Each time the predictor predicts $p_t = \frac{T}{2t}$ for $t > \frac{T}{2}$, the prediction can be swapped to the realization of state $\theta_t = 0$ to improve the score. The swap regret is calculated by

$$\text{SWAP}_\mu = \sum_{t>\frac{T}{2}} \text{VBREG}\left(\frac{T}{2t}, 0\right) = \sum_{t>\frac{T}{2}} \mathbb{I}\left[\frac{T}{2t} \geq \mu\right] = \frac{1}{2}\left(\frac{1}{\mu} - 1\right) T = \Theta(T),$$

implying $\text{MSR} = \Omega(T)$. □

5 Minimizing Maximum Swap Regret

In this section, we present our online binary prediction algorithm (Algorithm 1) and prove that it achieves the following low MSR guarantee:

Theorem 5.1. *For $T \geq 2$, Algorithm 1 runs in time polynomial in T and makes predictions $\boldsymbol{p} = (p_1, \ldots, p_T)$ satisfying*

$$\mathbf{E}\left[\mathrm{MSR}(\boldsymbol{p}, \boldsymbol{\theta})\right] \leq O(\sqrt{T} \log T).$$

Here, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_T)$ is the sequence of realized states chosen by any adversary in the online binary prediction setting (see Section 2.1), and the expectation is over the randomness of the algorithm.

We design Algorithm 1 such that it makes predictions in a finite set $Q := \{q_1, \ldots, q_m\} \subseteq [0, 1]$, where $q_i = i/m$ for each $i = 1, \ldots, m$. Later we will pick the optimal choice of $m \approx \sqrt{T}/\log T$. We view each q_i as a bucket, so the prediction p_t made by Algorithm 1 in each round t falls into one of the m buckets $q_1, \ldots, q_m$. We use n_i to denote the number of predictions in bucket i (see (2)), and use $\widehat{q}_i$ to denote the average value of the realized states corresponding to the n_i predictions (see (3)). We define $G_i := n_i|q_i - \widehat{q}_i|$ as the *bias* from bucket i.

In Section 5.1, we prove a key technical lemma which allows us to attribute the MSR to the bias G_i from each bucket. We then present Algorithm 1 in Section 5.2 and complete the proof of Theorem 5.1.

5.1 Attributing MSR to Bucket-wise Biases

We establish our key technical lemma that allows us to upper bound MSR using the bucket-wise biases.

Lemma 5.2 (Formal and generalized version of Lemma 1.1)**.** *Let $T, m \geq 2$ be positive integers. Define $Q = \{q_1, \ldots, q_m\} \subseteq [0, 1]$ where $q_i = i/m$ for every $i = 1, \ldots, m$. Given a sequence of predictions $\boldsymbol{p} = (p_1, \ldots, p_T) \in Q^T$ and realized states $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_T) \in \{0, 1\}^T$, define n_i and $\widehat{q}_i$ as in (2) and (3). Define $G_i := n_i|\widehat{q}_i - q_i|$. For $\alpha, \beta \geq 0$, define maximum deviation $\mathcal{D}$:*

$$\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta}) := \max_{1 \leq i \leq m} \{[G_i - \alpha\sqrt{n_i} - \beta n_i]_+\}, \quad \text{where } [x]_+ := \max(x, 0). \tag{19}$$

Then

$$\mathrm{MSR}(\boldsymbol{p}, \boldsymbol{\theta}) \leq 4m\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta}) + 4\alpha\sqrt{T} + 4\beta T + O(\alpha^2 m \log m).$$

Our proof of Lemma 5.2 relies on the following helper lemma which controls the swap regret w.r.t. a single V-shaped Bregman divergence.

Lemma 5.3. *Let T, m be positive integers. Define $Q = \{q_1, \ldots, q_m\} \subseteq [0, 1]$ where $q_i = i/m$ for every $i = 1, \ldots, m$. Given a sequence of predictions $\boldsymbol{p} = (p_1, \ldots, p_T) \in Q^T$ and realized states $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_T) \in \{0, 1\}^T$, define n_i and $\widehat{q}_i$ as in (2) and (3). Define $G_i := n_i|\widehat{q}_i - q_i|$. Fix a V-Bregman divergence with kink μ, the prediction swap regret is bounded by*

$$\textsc{Swap}_\mu(\boldsymbol{p}, \boldsymbol{\theta}) \leq 2 \sum_{i \in [m]} \left(G_i - n_i|q_i - \mu|\right)_+.$$

Proof.

$$\textsc{Swap}_\mu(\boldsymbol{p}, \boldsymbol{\theta}) = \sum_{i \in [m]} n_i \textsc{VBreg}_\mu(q_i, \widehat{q}_i)$$

$$
\begin{aligned}
&= \sum_{i\in[m]} n_i \frac{|\widehat{q}_i - \mu|}{\max\{1-\mu,\mu\}} \left(\mathbb{I}\left[q_i < \mu < \widehat{q}_i\right] + \mathbb{I}\left[q_i > \mu > \widehat{q}_i\right]\right) && \text{(by (14))}\\
&\le \sum_{i\in[m]} n_i \frac{(|\widehat{q}_i - q_i| - |\mu - q_i|)_+}{\max\{1-\mu,\mu\}}\\
&\le 2\sum_{i\in[m]} (G_i - n_i|q_i-\mu|)_+ . && \square
\end{aligned}
$$

Proof of Lemma 5.2. By Theorem 3.10, it suffices to prove

$$
\textsc{VSwap}(\boldsymbol{p},\boldsymbol{\theta}) \le 2m\mathcal{D}(\boldsymbol{p},\boldsymbol{\theta}) + 2\alpha\sqrt{T} + 2\beta T + O(\alpha^2 m\log m). \tag{20}
$$

For any $\mu\in[0,1]$, by Lemma 5.3,

$$
\begin{aligned}
\textsc{Swap}_\mu(\boldsymbol{p},\boldsymbol{\theta}) &\le 2\sum_{i=1}^m (G_i - n_i|q_i-\mu|)_+\\
&\le 2\sum_{i=1}^m (\mathcal{D}(\boldsymbol{p},\boldsymbol{\theta}) + \alpha\sqrt{n_i} + \beta n_i - n_i|q_i-\mu|)_+\\
&\le 2\sum_{i=1}^m (\mathcal{D}(\boldsymbol{p},\boldsymbol{\theta}) + \beta n_i + (\alpha\sqrt{n_i} - n_i|q_i-\mu|)_+) \quad \text{(because } \mathcal{D}(\boldsymbol{p},\boldsymbol{\theta})\ge 0 \text{ and } \beta\ge 0)\\
&= 2m\mathcal{D}(\boldsymbol{p},\boldsymbol{\theta}) + 2\beta T + 2\sum_{i=1}^m (\alpha\sqrt{n_i} - n_i|q_i-\mu|)_+ . \qquad (21)
\end{aligned}
$$

Let us re-arrange $q_1,\dots,q_m$ in non-decreasing order of $|q_i-\mu|$. That is, we choose a bijection τ from $\{1,\dots,m\}$ to itself such that $|q_{\tau(i)}-\mu|$ is a non-decreasing function of i. When $i=1$, we use the following trivial upper bound:

$$
(\alpha\sqrt{n_{\tau(1)}} - n_{\tau(1)}|q_{\tau(1)}-\mu|)_+ \le \alpha\sqrt{n_{\tau(1)}} \le \alpha\sqrt{T}. \tag{22}
$$

When $i>1$, we have $|q_{\tau(i)}-\mu| \ge \Omega(i/m)$, and thus

$$
(\alpha\sqrt{n_{\tau(i)}} - n_{\tau(i)}|q_{\tau(i)}-\mu|)_+ \le \frac{\alpha^2}{4|q_{\tau(i)}-\mu|} = O(\alpha^2 m/i). \tag{23}
$$

Plugging (22) and (23) into (21), we get

$$
\begin{aligned}
\textsc{Swap}_\mu(\boldsymbol{p},\boldsymbol{\theta}) &\le 2m\mathcal{D}(\boldsymbol{p},\boldsymbol{\theta}) + 2\beta T + 2\alpha\sqrt{T} + O\left(\alpha^2 m\sum_{i=2}^m \frac{1}{i}\right)\\
&\le 2m\mathcal{D}(\boldsymbol{p},\boldsymbol{\theta}) + 2\beta T + 2\alpha\sqrt{T} + O(\alpha^2 m\log m).
\end{aligned}
$$

This implies (20), as desired. $\square$

5.2 Efficient MSR Minimization Algorithm

Given Lemma 5.2, we can establish the low MSR guarantee in Theorem 5.1 by designing an algorithm (Algorithm 1) that minimizes $\mathcal{D}(\boldsymbol{p},\boldsymbol{\theta})$. Specifically, for parameters $m\approx\sqrt{T}/\log T, \beta = 1/m, \alpha\approx\sqrt{\log T}$, we show that Algorithm 1 achieves $\mathbf{E}\left[\mathcal{D}(\boldsymbol{p},\boldsymbol{\theta})\right] = O(\log T)$ (Lemma 5.6). Our

design of Algorithm 1 largely follows the ideas from Noarov et al. (2023), but we make small but important refinements to obtain a stronger guarantee as needed to prove Theorem 5.1 (see Remark 5.7).

In Algorithm 1, we partition the interval $[0,1]$ into m sub-intervals $I_1, \ldots, I_m$ where

$$I_1 = [0, 1/m], I_2 = (1/m, 2/m], \ldots, I_m = ((m-1)/m, 1]. \tag{24}$$

In each round $t = 1, \ldots, T$, Algorithm 1 first computes a prediction $\widetilde{p}_t \in [0,1]$ and then outputs a discretized prediction p_t via rounding. Specifically, $Q = \{q_1, \ldots, q_m\}$ is the discretized prediction space, where $q_i = i/m$ for every $i = 1, \ldots, m$. The prediction $\widetilde{p}_t$ belongs to an interval I_i for a unique index $i = 1, \ldots, m$, and the corresponding discretized prediction is $p_t = q_i \in Q$. We use n_i to denote the number of rounds t in which $p_t = q_i$, or equivalently, $\widetilde{p}_t \in I_i$:

$$n_i := \sum_{t=1}^{T} \mathbb{I}\left[p_t = q_i\right] = \sum_{t=1}^{T} \mathbb{I}\left[\widetilde{p}_t \in I_i\right].$$

For each $i = 1, \ldots, m$, and $\sigma = \pm 1$, we define

$$l_{i,\sigma}(p, \theta) := \sigma \mathbb{I}\left[p \in I_i\right](p - \theta) \quad \text{for every prediction } p \in [0,1] \text{ and state } \theta \in \{0,1\}. \tag{25}$$

Algorithm 1 calls an expert regret minimization oracle $\mathcal{A}$ from Chen et al. (2021). Here we imagine $2m+1$ *experts*: one expert for each pair $(i, \sigma) \in [m] \times \{\pm 1\}$ and one extra auxiliary expert. In each round t, the oracle $\mathcal{A}$ computes a distribution over the experts represented by values $w_{t,i,\sigma} \geq 0$, where for each pair (i, σ), the value $w_{t,i,\sigma} \geq 0$ is the probability mass on the expert corresponding to (i, σ). Thus, the probability mass on the auxiliary expert is $1 - \sum_{(i,\sigma)} w_{t,i,\sigma} \geq 0$.

The distribution computed by $\mathcal{A}$ in each round t is based on the *gains* $l_{t',i,\sigma} \in [-1, 1]$ received by each expert (i, σ) at Step 7 in each previous round $t' < t$. We set the gain of the auxiliary expert to always be zero. The work of Chen et al. (2021) shows a construction of the oracle with the following property:

Lemma 5.4 (Chen et al. (2021), applied to Algorithm 1)**.** *For some absolute constant $C > 0$, there exists an expert regret minimization oracle $\mathcal{A}$ for step 1 of Algorithm 1 with the following properties. Assume $T, m \geq 2$. In each round $t = 1, \ldots, T$, the oracle computes $w_{t,i,\sigma} \geq 0$ for every $i \in [m]$ and $\sigma = \pm 1$ in time $\mathsf{poly}(m)$ such that*

$$\sum_{i \in [m], \sigma = \pm 1} w_{t,i,\sigma} \leq 1,$$

$$-\sum_{t=1}^{T} \sum_{i \in [m], \sigma = \pm 1} w_{t,i,\sigma} l_{t,i,\sigma} \leq C \log(mT), \tag{26}$$

and for every $i = 1, \ldots, m$ and $\sigma = \pm 1$,

$$\sum_{t=1}^{T} l_{t,i,\sigma} - \sum_{t=1}^{T} \sum_{i' \in [m], \sigma' = \pm 1} w_{t,i',\sigma'} l_{t,i',\sigma'} \leq C \left(\log(mT) + \sqrt{n_i \log(mT)}\right). \tag{27}$$

For each expert (i, σ), the guarantee (27) is stronger than more standard guarantees for the experts problem in that the right hand side of (27) has an $\sqrt{n_i}$ dependence rather than an $\sqrt{T}$ dependence. For the auxiliary expert, we get the guarantee (26), which can be viewed as a special form of (27) with $l_{t,i,\sigma} = 0$ and $n_i = 0$.

The following lemma shows that Step 3 of Algorithm 1 can be computed efficiently:

Algorithm 1 Algorithm for MSR minimization.

Parameters: positive integers m, T; $\epsilon > 0$; discretized prediction space $Q = \{q_1, \ldots, q_m\}$ where $q_i = i/m$; intervals $I_1, \ldots, I_m$ partitioning $[0,1]$ as defined in (24); functions $l_{i,\sigma}$ as defined in (25).

for each round $t = 1, \ldots, T$ **do**

1. Compute expert weights $w_{t,i,\sigma}$ for every $i \in [m]$ and $\sigma = \pm 1$ using the expert regret minimization oracle $\mathcal{A}$ from Lemma 5.4.
2. For any distribution s over $[0,1]$, define

$$h_t(s) := \max_{\theta \in \{0,1\}} \mathbf{E}_{p \sim s} \left[\sum_{i \in [m], \sigma = \pm 1} w_{t,i,\sigma} l_{i,\sigma}(p, \theta) \right]. \tag{28}$$

3. Find distribution s_t such that $h_t(s_t) \le \epsilon$.
4. Draw $\widetilde{p}_t \in [0,1]$ from distribution s_t.
5. **Output** $p_t := q_i$, where i is the unique index in $\{1, \ldots, m\}$ satisfying $\widetilde{p}_t \in I_i$.
6. Receive the realized state $\theta_t \in \{0,1\}$.
7. Calculate expert gains $l_{t,i,\sigma}$ for every $i \in [m]$ and $\sigma = \pm 1$:

$$l_{t,i,\sigma} := l_{i,\sigma}(\widetilde{p}_t, \theta_t). \tag{29}$$

end for

Lemma 5.5 (Noarov et al. (2023))**.** *Let $\epsilon > 0$ be a parameter. At step 3 of Algorithm 1, a solution s_t satisfying $h_t(s_t) \le \epsilon$ always exists and can be computed in time $\mathsf{poly}(\epsilon^{-1})$.*

The existence of s_t in Lemma 5.5 can be proved using the minimax theorem. We refer the reader to Noarov et al. (2023) for a complete proof of (a more general version of) Lemma 5.5 which includes an efficient algorithm for computing s_t using the Follow-the-Perturbed-Leader approach.

We are now ready to prove that Algorithm 1 guarantees a small value of $\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta})$ in expectation:

Lemma 5.6. *Let $C > 0$ be the absolute constant from Lemma 5.4. Assume $m, T \ge 2$ and $\varepsilon = 1/T$ in Algorithm 1. Let $\boldsymbol{p} = (p_1, \ldots, p_T)$ be the predictions made by Algorithm 1 on the adversarially chosen states $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_T)$. Define $\mathcal{D}$ as in (19) for $\alpha = C\sqrt{\log(mT)}, \beta = 1/m$. Then,*

$$\mathbf{E}\left[\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta})\right] \le O(\log(mT)),$$

where the expectation is over the randomness of Algorithm 1.

Remark 5.7. *We define $\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta})$ in (19) as a maximum over the buckets i. Thus, Lemma 5.6 shows an upper bound on the "expectation of maximum", which is stronger than guarantee stated in Noarov et al. (2023) on the "maximum of expectation". In our proof below, we use a slightly more careful analysis than what is used by Noarov et al. (2023) to achieve the stronger guarantee.*

Proof of Lemma 5.6. For every $i = 1, \ldots, m$, by the definition of G_i in Lemma 5.2, we have

$$\begin{aligned} G_i = n_i |q_i - \widehat{q}_i| &= \left| \sum_{t=1}^{T} \mathbb{I}\left[p_t = q_i\right] (p_t - \theta_t) \right| \\ &\le \left| \sum_{t=1}^{T} \mathbb{I}\left[p_t = q_i\right] (\widetilde{p}_t - \theta_t) \right| + n_i/m \end{aligned}$$

$$= \max_{\sigma=\pm 1} \sum_{t=1}^{T} l_{i,\sigma}(\widetilde{p}_t, \theta_t) + n_i/m \qquad \text{(by (25))}$$

$$= \max_{\sigma=\pm 1} \sum_{t=1}^{T} l_{t,i,\sigma} + n_i/m. \qquad \text{(by (29))}$$

Therefore, by (27),

$$G_i - n_i/m - \alpha\sqrt{n_i} \le \left(\max_{\sigma=\pm 1} \sum_{t=1}^{T} l_{t,i,\sigma} \right) - \alpha\sqrt{n_i} \le \sum_{t=1}^{T} \sum_{i' \in [m], \sigma'=\pm 1} w_{t,i',\sigma'} l_{t,i',\sigma'} + C \log(mT).$$

By (26),

$$0 \le \sum_{t=1}^{T} \sum_{i \in [m], \sigma=\pm 1} w_{t,i,\sigma} l_{t,i,\sigma} + C \log(mT).$$

Combining the two inequalities above, we get

$$\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta}) \le \sum_{t=1}^{T} \sum_{i \in [m], \sigma=\pm 1} w_{t,i,\sigma} l_{t,i,\sigma} + C \log(mT). \tag{30}$$

By the definition of h_t in (28), the guarantee of $h_t(s_t) \le \epsilon$, and the fact that $\widetilde{p}_t$ is drawn from s_t, we get

$$\mathbf{E}\left[\sum_{t=1}^{T} \sum_{i \in [m], \sigma=\pm 1} w_{t,i,\sigma} l_{t,i,\sigma} \right] \le \sum_{t=1}^{T} \epsilon \le 1. \tag{31}$$

Combining (30) and (31), we get

$$\mathbf{E}\left[\mathcal{D}(\widetilde{\boldsymbol{p}}, \boldsymbol{\theta})\right] \le O(\log(mT)). \qquad \square$$

We now complete the proof of our main theorem.

Proof of Theorem 5.1. We choose $\epsilon = 1/T$ in Algorithm 1 and set $m = \Theta(\sqrt{T}/\log T)$. Following the setting of Lemma 5.6, we define $\mathcal{D}$ as in (19) for $\alpha = C\sqrt{\log(mT)}, \beta = 1/m$.

We have

$$\begin{aligned}
\mathbf{E}\left[\mathrm{MSR}(\boldsymbol{p}, \boldsymbol{\theta})\right] &\le 4m\mathbf{E}\left[\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta})\right] + 4\alpha\sqrt{T} + 4\beta T + O(\alpha^2 m \log m) && \text{(by Lemma 5.2)} \\
&\le O(m \log(mT)) + 4\alpha\sqrt{T} + 4\beta T + O(\alpha^2 m \log m) && \text{(by Lemma 5.6)} \\
&= O(m \log T) + O(\sqrt{T \log T}) + O(\sqrt{T} \log T) + O(\sqrt{T} \log T) \\
&= O(\sqrt{T} \log T).
\end{aligned}$$

The running time guarantee of Algorithm 1 follows from Lemmas 5.4 and 5.5. $\square$

6 The Economic Utility of Calibration: Why Swap Regret?

In this section, we discuss the economic utility of calibration in decision making and provide a justification of our maximum swap regret. We turn to the batch setting rather than the online adversarial setting.

It has been argued that calibration alone is not a metric for good prediction quality. Consider an example, where the state is drawn uniformly with $\Pr[\theta = 0] = \Pr[\theta = 1] = 1/2$, with the following three predictors.

(a) the perfect predictor always predicts $p_a = \theta$.

(b) the uninformative predictor always predicts $p_b = 1/2$.

(c) the miscalibrated predictor predicts $p_c = 1$ if $\theta = 1$ and $p_c = 0.6$ if $\theta = 0$.

Clearly, if the agent follows the prediction, predictor (a) yields a higher expected utility than both predictor (b) and (c). However, when it comes to the comparison between predictor (b) and (c), it is unclear which one is better. For example, suppose the agent faces a binary decision problem such that their optimal decision rule is to take different actions at threshold $p = 0.8$. Predictor (c) guarantees an expected payoff the same as (a) and strictly better than (b). As another example, suppose, instead, the agent has a binary decision problem with optimal threshold at 0.55. Then (c) gives an expected payoff strictly worst than (b).

From our perspective, the *calibration error* and the *informativeness* of a prediction are two orthogonal metrics of the prediction quality. Calibration may fail but the prediction can still be informative. For example, if the agent has been working with predictor (c) in the long run, the agent should be able to figure out a prediction of 0.6 implies the state $\theta = 0$. In such a case, the goal of calibration fails, because the decision making side has to learn to interpret the meaning of a prediction. However, predictor (c) is informative, conveying the same separation of states as the perfect predictor (a).

The informativeness of a prediction is the payoff upperbound under a stochastic setting. The economics literature defines the informativeness of a signal as the expected payoff when the agent best responds to the bayesian posterior on that signal. We treat the prediction as a signal correlated with the state. Suppose the prediction $p \in \Delta(\Theta)$ and the state $\theta \in \Theta$ are drawn from a joint distribution $\Delta(\Delta(\Theta) \times \Theta)$. This could happen when there exists a joint distribution $\Delta(\mathcal{X} \times \Theta)$ over some feature space $\mathcal{X}$ and the state space, and the prediction is a function of the features in $\mathcal{X}$. The informativeness of the prediction is the expected payoff from this bayesian posterior $\Pr[\theta|p]$. This expected payoff from bayesian posterior serves as an upperbound of expected payoff if the agent is only given this prediction.

Once informativeness is separated out, the economic utility of calibration is clear. The average prediction swap regret measures the loss in payoff from the upperbound of informativeness. Under a stochastic setting, the prediction is perfectly calibrated if and only if the average prediction swap regret $\lim_{T\to\infty} \frac{1}{T}\textsc{Swap}_S$ is 0 for all decision problems with scoring rule $S \in [0, 1]$. Swapping each prediction to $p = \Pr[\theta|p]$ calibrates the prediction, which corresponds to the modification rule in prediction swap regret. To this end, calibration recovers the value of information from the prediction and achieves its informativeness upperbound on expected payoff.

References

Anagnostides, Ioannis, Constantinos Daskalakis, Gabriele Farina, Maxwell Fishelson, Noah Golowich, and Tuomas Sandholm (2022) "Near-optimal no-regret learning for correlated equilibria in multi-player general-sum games," in *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, 736–749. 8

Arora, Sanjeev, Elad Hazan, and Satyen Kale (2012) "The Multiplicative Weights Update Method: a Meta-Algorithm and Applications," *Theory of Computing*, 8 (6), 121–164, 10.4086/toc.2012.v008a006. 3

Arunachaleswaran, Eshwar Ram, Natalie Collina, Aaron Roth, and Mirah Shi (2024) "An Elementary Predictor Obtaining $2\sqrt{T}$ Distance to Calibration," *arXiv preprint arXiv:2402.11410*. 3, 8, 11

Blackwell, David (1951) "Comparison of experiments," in *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, 2, 93–103, University of California Press. 7

Błasiok, Jarosław, Parikshit Gopalan, Lunjia Hu, Adam Tauman Kalai, and Preetum Nakkiran (2024) "Loss Minimization Yields Multicalibration for Large Neural Networks," in Guruswami, Venkatesan ed. *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*, 287 of Leibniz International Proceedings in Informatics (LIPIcs), 17:1–17:21, Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 10.4230/LIPIcs.ITCS.2024.17. 7

Błasiok, Jarosław, Parikshit Gopalan, Lunjia Hu, and Preetum Nakkiran (2023a) "A unifying theory of distance from calibration," in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, 1727–1740. 7, 11

Błasiok, Jarosław, Parikshit Gopalan, Lunjia Hu, and Preetum Nakkiran (2023b) "When Does Optimizing a Proper Loss Yield Calibration?" in Oh, A., T. Neumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine eds. *Advances in Neural Information Processing Systems*, 36, 72071–72095: Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2023/file/e4165c96702bac5f4962b70f3cf2f136-Paper-Conference.pdf. 7

Blum, Avrim and Yishay Mansour (2007) "From External to Internal Regret," *Journal of Machine Learning Research*, 8 (47), 1307–1324, http://jmlr.org/papers/v8/blum07a.html. 2, 8

Chen, Liyu, Haipeng Luo, and Chen-Yu Wei (2021) "Impossible tuning made possible: A new expert algorithm and its applications," in *Conference on Learning Theory*, 1216–1259, PMLR. 24

Dagan, Yuval, Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich (2023) "From external to swap regret 2.0: An efficient reduction and oblivious adversary for large action spaces," *arXiv preprint arXiv:2310.19786*. 8

Dawid, A. P. (1982) "The Well-Calibrated Bayesian," *Journal of the American Statistical Association*, 77 (379), 605–610, 10.1080/01621459.1982.10477856. 1

Foster, Dean P and Sergiu Hart (2021) "Forecast hedging and calibration," *Journal of Political Economy*, 129 (12), 3447–3490. 2, 10

Foster, Dean P and Rakesh Vohra (1999) "Regret in the on-line decision problem," *Games and Economic Behavior*, 29 (1-2), 7–35. 11

Foster, Dean P and Rakesh V Vohra (1997) "Calibrated learning and correlated equilibrium," *Games and Economic Behavior*, 21 (1-2), 40–55. 2, 8

Foster, Dean P and Rakesh V Vohra (1998) "Asymptotic calibration," *Biometrika*, 85 (2), 379–390. 2, 3, 5, 8, 9, 10, 13

Frongillo, Rafael and Ian Kash (2014) "General truthfulness characterizations via convex analysis," in *Web and Internet Economics: 10th International Conference, WINE 2014, Beijing, China, December 14-17, 2014. Proceedings 10*, 354–370, Springer. 12

Garg, Sumegha, Christopher Jung, Omer Reingold, and Aaron Roth (2024) "Oracle Efficient Online Multicalibration and Omniprediction," in *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2725–2792, 10.1137/1.9781611977912.98. 8

Gopalan, Parikshit, Lunjia Hu, Michael P. Kim, Omer Reingold, and Udi Wieder (2023a) "Loss Minimization Through the Lens Of Outcome Indistinguishability," in Tauman Kalai, Yael ed. *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, 251 of Leibniz International Proceedings in Informatics (LIPIcs), 60:1–60:20, Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 10.4230/LIPIcs.ITCS.2023.60. 8

Gopalan, Parikshit, Adam Tauman Kalai, Omer Reingold, Vatsal Sharan, and Udi Wieder (2022) "Omnipredictors," in Braverman, Mark ed. *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, 215 of Leibniz International Proceedings in Informatics (LIPIcs), 79:1–79:21, Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 10.4230/LIPIcs.ITCS.2022.79. 8

Gopalan, Parikshit, Michael Kim, and Omer Reingold (2023b) "Swap Agnostic Learning, or Characterizing Omniprediction via Multicalibration," in Oh, A., T. Neumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine eds. *Advances in Neural Information Processing Systems*, 36, 39936–39956: Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2023/file/7d693203215325902ff9dbdd067a50ac-Paper-Conference.pdf. 8

Gopalan, Parikshit, Princewill Okoroafor, Prasad Raghavendra, Abhishek Shetty, and Mihir Singhal (2024) "Omnipredictors for Regression and the Approximate Rank of Convex Functions," *arXiv preprint arXiv:2401.14645*. 8

Hart, Sergiu (2022) "Calibrated Forecasts: The Minimax Proof," *arXiv preprint arXiv:2209.05863*. 2, 6, 10

Hart, Sergiu and Andreu Mas-Colell (2000) "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, 68 (5), 1127–1150. 8

Hart, Sergiu and Andreu Mas-Colell (2001) "A reinforcement procedure leading to correlated equilibrium," in *Economics Essays: A Festschrift for Werner Hildenbrand*, 181–200: Springer. 8

Hart, Sergiu and Andreu Mas-Colell (2013) *Simple adaptive strategies: from regret-matching to uncoupled dynamics*, 4: World Scientific. 2

Hartline, Jason D ⓡ Liren Shan ⓡ Yingkai Li ⓡ Yifan Wu (2023) "Optimal scoring rules for multi-dimensional effort," in *The Thirty Sixth Annual Conference on Learning Theory*, 2624–2650, PMLR. 8

Hebert-Johnson, Ursula, Michael Kim, Omer Reingold, and Guy Rothblum (2018) "Multicalibration: Calibration for the (Computationally-Identifiable) Masses," in Dy, Jennifer and Andreas Krause eds. *Proceedings of the 35th International Conference on Machine Learning*, 80 of Proceedings of Machine Learning Research, 1939–1948: PMLR, 10–15 Jul, https://proceedings.mlr.press/v80/hebert-johnson18a.html. 8

Hu, Lunjia, Inbal Rachel Livni Navon, Omer Reingold, and Chutong Yang (2023) "Omnipredictors for Constrained Optimization," in Krause, Andreas, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett eds. *Proceedings of the 40th International Conference on Machine Learning*, 202 of Proceedings of Machine Learning Research, 13497–13527: PMLR, 23–29 Jul, https://proceedings.mlr.press/v202/hu23b.html. 8

Kakade, Sham M. and Dean P. Foster (2008) "Deterministic calibration and Nash equilibrium," *Journal of Computer and System Sciences*, 74 (1), 115–130, https://doi.org/10.1016/j.jcss.2007.04.017, Learning Theory 2004. 10

Kim, Michael P., Amirata Ghorbani, and James Zou (2019) "Multiaccuracy: Black-Box Post-Processing for Fairness in Classification," in *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '19, 247–254, New York, NY, USA: Association for Computing Machinery, 10.1145/3306618.3314287. 8

Kim, Michael P. and Juan C. Perdomo (2023) "Making Decisions Under Outcome Performativity," in Tauman Kalai, Yael ed. *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, 251 of Leibniz International Proceedings in Informatics (LIPIcs), 79:1–79:15, Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 10.4230/LIPIcs.ITCS.2023.79. 8

Kleinberg, Bobby, Renato Paes Leme, Jon Schneider, and Yifeng Teng (2023) "U-calibration: Forecasting for an unknown agent," in *The Thirty Sixth Annual Conference on Learning Theory*, 5143–5145, PMLR. 1, 2, 3, 4, 6, 8, 11, 13, 16, 18, 19, 21

Li, Yingkai ⓡ Jason D Hartline ⓡ Liren Shan ⓡ Yifan Wu (2022) "Optimization of scoring rules," in *Proceedings of the 23rd ACM Conference on Economics and Computation*, 988–989. 3, 5, 8, 9, 16, 18

McCarthy, John (1956) "Measures of the value of information," *Proceedings of the National Academy of Sciences of the United States of America*, 42 (9), 654. 14, 15

Neyman, Eric, Georgy Noarov, and S Matthew Weinberg (2021) "Binary scoring rules that incentivize precision," in *Proceedings of the 22nd ACM Conference on Economics and Computation*, 718–733. 8

Noarov, Georgy, Ramya Ramalingam, Aaron Roth, and Stephan Xie (2023) "High-Dimensional Unbiased Prediction for Sequential Decision Making," in *OPT 2023: Optimization for Machine Learning*, https://openreview.net/forum?id=P4j4l45NUq. 1, 3, 4, 6, 8, 24, 25

Peng, Binghui and Aviad Rubinstein (2023) "Fast swap regret minimization and applications to approximate correlated equilibria," *arXiv preprint arXiv:2310.19647*. 2, 8

Qiao, Mingda and Gregory Valiant (2021) "Stronger calibration lower bounds via sidestepping," in *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2021, 456–466, New York, NY, USA: Association for Computing Machinery, 10.1145/3406325.3451050. 1, 2, 3, 7, 8, 10

Qiao, Mingda and Letian Zheng (2024) "On the Distance from Calibration in Sequential Prediction," *arXiv preprint arXiv:2402.07458*. 3, 8, 11

Roth, Aaron (2022) "Uncertain: Modern topics in uncertainty estimation." 7, 10, 19

Roth, Aaron and Mirah Shi (2024) "Forecasting for Swap Regret for All Downstream Agents," *arXiv preprint arXiv:2402.08753*. 1, 3, 4, 8, 11

Savage, Leonard J (1971) "Elicitation of personal probabilities and expectations," *Journal of the American Statistical Association*, 66 (336), 783–801. 14, 15

A Minimax Proof for Minimizing MSR

We can prove the existence of an algorithm that achieves $\widetilde{O}(\sqrt{T})$ MSR via the minimax theorem. The minimax theorem demonstrates a remarkable difference between K_1 and MSR - simply by reporting the adversary's strategy truthfully, MSR matches the lowerbound.

In this setup, a predictor F makes a prediction $p_t \in [0,1]$ at each time step $t = 1, 2, \ldots$, and an adversary A picks an outcome $\theta_t \in \{0,1\}$. Both p_t and θ_t are chosen based on the history $h_{t-1} = (p_1, \theta_1, p_2, \theta_2, \ldots, p_{t-1}, \theta_{t-1})$, but these two quantities at the same time step t cannot depend on each other. The goal of the predictor is to minimize $\mathrm{MSR}(h_T)$.

We can identify a predictor F by its strategy at each time step. That is, we can write $F = (F_1, \ldots, F_T)$, where F_t is a function mapping from h_{t-1} to p_t. Similarly, we can identify the adversary A by its strategy A_t at each time step t, writing A as $(A_1, \ldots, A_T)$. Given the strategies F and A of both players, we use $h_{F,A}$ to denote the history $h_T = (p_1, \theta_1, \ldots, p_T, \theta_T)$ generated by executing these strategies.

We will allow the predictor to randomize and play a mixed strategy which can be identified as a probability distribution $\mathcal{F}$ over all predictor strategies F.

Theorem A.1 (Existence)**.** *There exists a mixed predictor strategy $\mathcal{F}$ such that*

$$\max_A \mathbb{E}_{F\sim\mathcal{F}}[\mathrm{MSR}(h_{F,A})] = O(\log(T)\sqrt{T}),$$

where the maximum is over all strategies A of the adversary.

In fact, we will prove a stronger version of Theorem A.1, where we restrict the predictor to make predictions p_t in a discretized space $Q = \{q_i = \frac{i}{m}\}_{i\in[m]} \subseteq [0,1]$. We say a predictor strategy $F = (F_1, \ldots, F_T)$ is *discretized* if each F_t is a function mapping from history h_{t-1} to $p_t \in Q$. This discretization makes the (pure) strategy space of the predictor to be finite, allowing us to apply the minimax theorem.

We prove the following argument which implies Theorem A.1 immediately:

$$\min_{\mathcal{F}} \max_A \mathbb{E}_{F\sim\mathcal{F}}[\mathrm{MSR}(h_{F,A})] = O(\log(T)\sqrt{T}),$$

where the minimum is over a distribution $\mathcal{F}$ over discretized predictor strategies F, and the maximum is over all strategies A of the adversary.

By the minimax theorem,

$$\min_{\mathcal{F}} \max_A \mathbb{E}_{F\sim\mathcal{F}}[\mathrm{MSR}(h_{F,A})] = \max_{\mathcal{A}} \min_F \mathbb{E}_{A\sim\mathcal{A}}[\mathrm{MSR}(h_{F,A})].$$

Thus, it suffices to reverse the order of play and prove Lemma A.2.

Lemma A.2. *The following holds:*

$$\max_{\mathcal{A}} \min_F \mathbb{E}_{A\sim\mathcal{A}}[\mathrm{MSR}(h_{F,A})] \le O(\log(T)\sqrt{T}).$$

where the minimum is over a deterministic discretized predictor strategy F, and the maximum is over a distribution over strategies A of the adversary.

To show Lemma A.2, consider the truthful strategy of the predictor. At round t, let $\widetilde{p}_t$ denote the conditional expectation of the adversary's choice θ_t given past history. The predictor outputs the closest value p_t in Q to $\widetilde{p}_t$.

Lemma A.3. *Let $\boldsymbol{p} = (p_1, \ldots, p_T)$ and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_T)$ be the sequence of predictions and states generated when the predictor uses the truthful strategy above. For $\alpha \geq 0$ and $\beta = \frac{1}{m}$, define $\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta})$ as in Equation (19). We have*

$$\mathbf{E}\left[\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta})\right] \leq 2T^2 m \exp\left(-\alpha^2/2\right).$$

Proof of Lemma A.3. For bucket i, construct random variables $X_1^{(i)}, \ldots X_T^{(i)}$ such that $X_j^{(i)} = \theta_{t_j} - \widetilde{p}_{t_j}$, where t_j is the index of the j-th round in which the predictor predicts q_i. If the number n_i of rounds with prediction q_i is smaller than j, we define $X_j^{(i)} = 0$. For any $n \in [T]$, by Azuma's inequality,

$$\Pr\left[\left|\sum_{j=1}^{n} X_j^{(i)}\right| > \alpha\sqrt{n}\right] \leq 2\exp\left(\frac{-\alpha^2}{2}\right).$$

By the union bound,

$$\Pr\left[\exists n \in [T], i \in [m], \left|\sum_{j=1}^{n} X_j^{(i)}\right| > \alpha\sqrt{n}\right] \leq 2Tm\exp\left(\frac{-\alpha^2}{2}\right). \tag{32}$$

Therefore, with probability at least $1 - 2Tm\exp\left(\frac{-\alpha^2}{2}\right)$, for every $i \in [m]$, we have

$$\left|\sum_{j=1}^{n_i} X_j^{(i)}\right| \leq \alpha\sqrt{n_i}.$$

This implies

$$G_i = \left|\sum_{j=1}^{n_i}(p_{t_j} - \theta_{t_j})\right| \leq \left|\sum_{j=1}^{n_i}(\widetilde{p}_{t_j} - \theta_{t_j})\right| + n_i/m \leq \alpha\sqrt{n_i} + \beta n_i.$$

Thus, $\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta}) = 0$ with probability at least $1 - 2Tm\exp\left(\frac{-\alpha^2}{2}\right)$. Since $\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta}) \leq T$,

$$\mathbf{E}\left[\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta})\right] \leq 2T^2 m \exp\left(\frac{-\alpha^2}{2}\right).$$

□

Proof of Lemma A.2. Taking $\alpha = 2\sqrt{\log(Tm)}$ in Lemma A.3, we get

$$\mathbf{E}\left[\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta})\right] \leq 2.$$

We choose $m = \Theta(\sqrt{T}/\log(T))$ buckets and set $\beta = \frac{1}{m}$. By Lemma 5.2,

$$\begin{aligned}
\mathbf{E}\left[\mathrm{MSR}(\boldsymbol{p}, \boldsymbol{\theta})\right] &\leq 4m\mathbf{E}\left[\mathcal{D}(\boldsymbol{p}, \boldsymbol{\theta})\right] + 4\alpha\sqrt{T} + 4\beta T + O(\alpha^2 m \log m) \\
&\leq O(\sqrt{T}/\log(T)) + O(\sqrt{T\log(T)}) + O(\sqrt{T}\log T) + O(\sqrt{T}\log(T)) \\
&= O(\sqrt{T}\log T).
\end{aligned}$$

□