Proceedings of Machine Learning Research vol 336:1–26, 2026 39th Annual Conference on Learning Theory

Calibeating Made Simple

Yurong Chen YURONG.CHEN@INRIA.FR
Inria, École Normale Supérieure, PSL Research University

Zhiyi Huang ZHIYI@CS.HKU.HK
The University of Hong Kong

Michael I. Jordan JORDAN@CS.BERKELEY.EDU
Inria, École Normale Supérieure, PSL Research University, University of California, Berkeley

Haipeng Luo HAIPENGL@USC.EDU
University of Southern California

Editors: Steve Hanneke and Tor Lattimore

Abstract

We study calibeating, the problem of post-processing external forecasts online to minimize cumulative losses and match an informativeness-based benchmark. Unlike prior work, which analyzed calibeating for specific losses with specific arguments, we reduce calibeating to existing online learning techniques and obtain results for general proper losses. More concretely, we first show that calibeating is minimax-equivalent to regret minimization. This recovers the $O(\log T)$ calibeating rate of Foster and Hart (2023) for the Brier and log losses and its optimality, and yields new optimal calibeating rates for exp-concave losses and general bounded losses. Second, we prove that multi-calibeating is minimax-equivalent to the combination of calibeating and the classical expert problem. This yields new optimal multi-calibeating rates for exp-concave losses, including Brier and log losses, and general bounded losses. Finally, we obtain new bounds for achieving calibeating and calibration simultaneously for the Brier loss. For binary predictions, our result gives the first calibrated algorithm that at the same time also achieves the optimal $O(\log T)$ calibeating rate.

Keywords: Calibeating, online calibration, probabilistic forecasting, proper scoring rules, regret minimization, online learning

1. Introduction

Calibration has attracted growing attention in recent years as a desideratum for probabilistic prediction, motivated by the need to produce reliable probabilities for downstream decision-making (Guo et al., 2017). Despite its appeal as a benchmark for reliability, however, calibration is not necessarily a meaningful test of forecasting expertise. For example, online calibration can be achieved by randomized strategies without any knowledge of the data-generating process (Foster and Vohra, 1998). Hence, calibration alone cannot distinguish true expertise from uninformative procedures.

To quantify and preserve forecasting expertise, Foster and Hart (2023) introduced *calibeating* in a post-processing setting. In this setting, an external forecaster (e.g., a machine learning model) outputs a probabilistic forecast at each round, and then the learner produces its own forecast based on it. It is known that for proper losses such as the Brier and log losses, the cumulative score can be decomposed into a calibration term, which measures the reliability, and a refinement term, which measures the informativeness and skill. This motivates the question of whether one can improve reliability without sacrificing skill. Calibeating formalizes this goal by requiring the learner's loss to be as small as the external forecaster's refinement score (in other words, to "beat" the forecaster by its calibration error).

Existing work (Foster and Hart, 2023; Lee et al., 2022) establishes online calibeating guarantees for the Brier and log losses, and studies extensions such as beating multiple forecasters (multi-calibeating) and imposing simultaneous calibration constraints. These results rely on loss-specialized analyses. More broadly, the fundamental statistical difficulty of calibeating and its relationship to standard online-learning problems have remained unclear, leaving open whether known bounds are optimal or how they generalize beyond the Brier and log losses.

1.1. Our Results

We study calibeating from an online-learning perspective. Rather than analyzing different losses on a case-by-case basis, we identify simple reductions from (multi-)calibeating to standard online-learning primitives. This yields a "plug-and-play" analysis: by instantiating the reductions with classical online-learning algorithms, we obtain general upper and lower bounds in a modular way.

Calibeating $=$ no-regret learning (Section 3). We prove that calibeating is minimax-equivalent to regret minimization. Theorem 4 gives a reduction that turns any no-regret learner with regret bound $\alpha(T)$ into a calibeating algorithm with a corresponding bound that scales with $|Q|$, the number of distinct external forecast values over T rounds. The reduction exploits the fact that the refinement benchmark decomposes across distinct forecast values, allowing one to treat each corresponding subsequence independently. Instantiating this reduction recovers the $O(|Q| \log T)$ guarantees [1] for the Brier and log losses from Foster and Hart (2023) and extends them to general exp-concave losses (Corollary 7). We also obtain an $O(\sqrt{|Q|T})$ bound for general bounded proper losses (Corollary 5). Conversely, Theorem 8 provides a matching lower bound that completes the minimax-equivalence.

Multi-calibeating $=$ calibeating $+$ expert problem (Section 4). Next, we present (in Theorem 11) a simple decomposition of multi-calibeating into calibeating and the expert problem: run a separate calibeating subroutine for each forecaster to produce candidate predictions, then aggregate them with an expert algorithm. The resulting multi-calibeating guarantee is the sum of the calibeating bound and the expert regret bound. For exp-concave losses, we obtain a logarithmic bound of $O(\log N + |Q^{(n)}| \log T)$ (Corollary 13), where N is the number of forecasters and $Q^{(n)}$ is the set of distinct forecasts produced by forecaster n. This improves exponentially over the polynomial dependence on N in Foster and Hart (2023) and the polynomial dependence on T in Lee et al. (2022). We complement this with a lower-bound reduction (Theorem 14), showing that multi-calibeating inherits hardness from both the expert problem and the calibeating problem. This yields matching lower bounds for Brier and log losses and shows the tightness of our results (Corollary 16).

Simultaneous (multi-)calibeating and calibration for Brier loss (Section 5). Finally, we provide new bounds for achieving calibeating and calibration simultaneously. We propose a meta-algorithm that tracks an arbitrary reference algorithm while ensuring calibration (Theorem 17). The construction employs two existing online learning primitives: the reduction by Blum and Mansour (2007) to enforce calibration via the calibration-swap-regret connection, and a two-expert algorithm by Sani et al. (2014) to aggregate the predictions from the Blum–Mansour (BM) reduction and the

1. We omit the polynomial dependence on the number of outcomes for presentation clarity and refer the detailed bound to the theorem statements.

Table 1: Comparison of prior and our guarantees in N, T, K (the number of outcomes), and $|Q|$ (we assume $|Q^{(n)}| = |Q|$ for simplicity). For simultaneous calibeating and calibration, the first rate is for calibeating and the second for calibration. We omit polynomial dependence on K for presentation clarity, and $\tilde{O}$ omits logarithmic dependence on T. The simultaneous results of Foster and Hart (2023) are only for calibeating (but not multi-calibeating), so we only show results for calibeating for comparison. The results of Lee et al. (2022) are only for binary outcomes.

Setting	Loss class	Prior work	**This paper**	
Calibeating	Exp-concave	–	$O(\|Q\| \log T)$	(Cor. 7, 10)
	- Brier	$\Theta(\|Q\| \log T)$ [1]	$\Theta(\|Q\| \log T)$	
	- Log	$O(\|Q\| \log T)$ [1]	$\Theta(\|Q\| \log T)$	
	Bounded	–	$\Theta(\sqrt{\|Q\|T})$	(Cor. 5, 9)
Multi-Calibeating	Exp-concave	–	$\Theta(\log N + \|Q\| \log T)$	(Cor. 13, 16)
	- Brier	$O((N + \|Q\|) \log T)$ [1] $O(\sqrt{NT} + \|Q\| \log T)$ [1] $\tilde{O}(\sqrt{\|Q\|}(\log N)^{\frac{1}{4}} T^{\frac{3}{4}})$ [2]		
	Bounded	–	$\Theta(\sqrt{T \log N} + \sqrt{\|Q\|T})$	(Cor. 12, 15)
Calibeating & Calibration	Brier			
	- binary	$\tilde{O}(\|Q\|^{\frac{2}{3}} T^{\frac{1}{3}}), \tilde{O}(\|Q\|^{\frac{2}{3}} T^{\frac{1}{3}})$ [1]	$O(\|Q\| \log T), \tilde{O}(\sqrt{T})$	(Cor. 18)
	- K-class	$\tilde{O}(\|Q\|^{\frac{2}{K+1}} T^{\frac{K-1}{K+1}}), \tilde{O}(\|Q\|^{\frac{2}{K+1}} T^{\frac{K-1}{K+1}})$ [1]	$\tilde{O}(\|Q\| + T^{\frac{K-1}{K+1}}), \tilde{O}(T^{\frac{K-1}{K+1}})$	(Cor. 19)

[1] Foster and Hart 2023, [2] Lee et al. 2022

reference algorithm. Instantiating the reference with the (multi-)calibeating algorithms from the previous sections, for the Brier loss, we obtain for the binary case the optimal logarithmic (multi-)calibeating rate of $O(\log N + |Q^{(n)}| \log T)$ while ensuring a sublinear ℓ_2-calibration error of order $\tilde{O}(\sqrt{T})$ (Corollary 18). This improves the polynomial T-dependence on the calibeating side in Foster and Hart (2023) and improves both sides compared to Lee et al. (2022). For multi-class outcomes, we derive explicit tradeoffs between (multi-)calibeating and calibration (Corollary 19). In particular, at one extreme, we recover the known dependence on T for calibration (Foster and Hart, 2023; Fishelson et al., 2025) while dropping the $|Q^{(n)}|$ dependence in Foster and Hart (2023). For a summary of our results and comparisons with prior work, see Table 1.

1.2. Related Work

Calibeating. In the seminal work proposing calibeating, Foster and Hart (2023) give online guarantees for the Brier and log losses via a bin-wise estimation viewpoint. They also study extensions to multiple forecasters and to simultaneous calibeating and calibration. Lee et al. (2022) formulate simultaneous multi-calibration and multi-calibeating as an online multi-objective optimization problem, achieving favorable dependence on the number of external forecasters but suboptimal dependence on the time horizon. In comparison, our results are obtained via reductions that connect calibeating to standard online-learning problems. Finally, as an application, Gupta and Ramdas (2023) apply calibeating as a robustness layer on top of online Platt scaling to guarantee adversarial calibration in binary classification while preserving predictive performance.

Online recalibration. Online recalibration is studied in the same post-processing setting as calibeating, but it benchmarks performance by proper-loss regret rather than refinement. The goal is to achieve small regret relative to the external forecaster while simultaneously ensuring calibrated predictions (Marx et al., 2025; Deshpande et al., 2024). For binary classification, Kuleshov and Ermon (2017) provide adversarial online guarantees, and Okoroafor et al. (2024) obtain improved bounds and explicit regret versus ℓ_1-calibration tradeoffs via Blackwell approachability for strictly proper losses. These tradeoffs yield sublinear but typically polynomial-in-T rates. While we focus on the Brier loss in our simultaneous guarantee, we target the stronger refinement benchmark and achieve logarithmic-in-T rates while still ensuring sublinear ℓ_2-calibration.

Calibration and proper scoring loss. Proper scoring losses admit classical decompositions into a reliability (calibration) term and an informativeness (refinement) term (Dawid, 2006; Sanders, 1963; Bröcker, 2009). In this spirit, ℓ_2-calibration (Foster and Vohra, 1998) and KL-calibration (Luo et al., 2025) can be viewed as online analogues of the calibration term for the Brier loss and the log loss, respectively; more generally, this motivates defining online calibration measures compatible with arbitrary proper scoring losses. Several calibration notions, including ℓ_2-calibration (Fishelson et al., 2025) and KL-calibration (Luo et al., 2025), have been shown to be equivalent to swap-regret objectives. We exploit this connection in our simultaneous guarantees by enforcing (pseudo-)swap regret via the BM reduction (Blum and Mansour, 2007).

2. Model

We consider an online prediction problem over a finite outcome space with N external forecasts. Let $K \geq 2$ be the number of possible outcomes, and $\Delta_K := \{p \in \mathbb{R}^K_{\geq 0} : \sum_{k=1}^K p_k = 1\}$ be the probability simplex. We let $[n]$ denote the set $\{1, \dots, n\}$ for any positive integer n. The outcome space is denoted by $\mathcal{E} := \{e_i : i \in [K]\} \subseteq \Delta_K$, where e_i is the i-th standard basis vector.

The interaction proceeds for T rounds. At each round $t \in [T]$, the learner first observes N external forecasts, $q_t^{(n)} \in \Delta_K$, $n \in [N]$, and makes its own prediction $p_t \in \Delta_K$. The outcome $y_t \in \mathcal{E}$ is then revealed, and the learner incurs loss $\ell(p_t, y_t)$. For simplicity, we assume that $q_{1:T} := (q_t)_{t=1}^T$ and $y_{1:T} := (y_t)_{t=1}^T$ are generated by an oblivious adversary, i.e., they are decided at time $t = 0$ with complete knowledge of the learner's algorithm (but not its random bits).

Throughout, we consider a proper scoring loss $\ell : \Delta_K \times \mathcal{E} \to \mathbb{R}$, i.e., losses such that for any $q \in \Delta_K$, $q \in \arg\min_{p \in \Delta_K} \mathbb{E}_{y \sim q}[\ell(p, y)]$. We write $\ell(p, q) := \mathbb{E}_{y \sim q}[\ell(p, y)]$. Let $\mathbf{1}\{\cdot\}$ denote the indicator function, which equals one if the condition holds and zero otherwise. Given a prediction sequence $p_{1:T}$ and outcome sequence $y_{1:T}$, for any $p \in \Delta_K$, denote the number of times the learner predicts p as $n_T(p) := \sum_{t=1}^T \mathbf{1}\{p_t = p\}$, and the empirical outcome distribution conditioned on prediction p as $\rho_T^p(y) := \frac{1}{n_T(p)} \sum_{t=1}^T \mathbf{1}\{p_t = p, y_t = y\}$ for $y \in \mathcal{E}$, whenever $n_T(p) > 0$. With these definitions, the cumulative loss, refinement score, and calibration error are defined as follows.

Definition 1 *The* cumulative loss *of predictions $p_{1:T}$ under outcomes $y_{1:T}$ is*

$$L_T(p_{1:T}, y_{1:T}) := \sum_{t=1}^T \ell(p_t, y_t) .$$

The refinement score *is*

$$R_T(p_{1:T}, y_{1:T}) \coloneqq \sum_p n_T(p)\ell(\rho_T^p, \rho_T^p) = \sum_p \min_{q\in\Delta_K} \sum_{t:p_t=p} \ell(q, y_t) \, .$$

Finally, the calibration error *is*

$$K_T(p_{1:T}, y_{1:T}) \coloneqq L_T(p_{1:T}, y_{1:T}) - R_T(p_{1:T}, y_{1:T}) \, .$$

By construction, $L_T = R_T + K_T$ and $K_T \geq 0$. Moreover, K_T coincides with the full-swap-regret notion of Fishelson et al. (2025), while R_T corresponds to the best-in-hindsight swap-regret benchmark. Indeed, for each prediction p, the refinement term equals the loss of the best constant predictor over rounds with $p_t = p$. Thus, the refinement score measures the informativeness of the forecasts: sequences that induce finer bins with lower within-bin variability achieve smaller refinement. In contrast, the calibration error measures within-bin reliability, i.e., how close the issued prediction p is to the empirical conditional distribution ρ_T^p on the corresponding subsequence.

Proper scoring losses admit a classic decomposition into terms measuring the informativeness (or refinement) of forecasts and their reliability (or calibration) in a probabilistic setting; see, e.g., Bröcker (2009); Dawid (2006). Definition 1 can be seen as the empirical counterparts of these quantities.

Example 1 *For* Brier loss $\ell(p, y) = \|p - y\|_2^2$, *the refinement score equals the weighted sum of within-bin variances, and the calibration error becomes the* ℓ_2*-calibration (Foster and Vohra, 1998),*

$$R_T(p_{1:T}, y_{1:T}) = \sum_p n_T(p) \sum_{t:p_t=p} \frac{1}{n_T(p)} \|\rho_T^p - y_t\|_2^2 \, , \quad K_T(p_{1:T}, y_{1:T}) = \sum_p n_T(p)\|p - \rho_T^p\|_2^2 \, .$$

Example 2 *Denote the Shannon entropy under distribution* p *to be* $H(p) = -\sum_k p_k \log p_k$*. For* log loss $\ell(p, y) = -\sum_{k=1}^K y_k \log p_k$*, the refinement score equals the weighted sum of the Shannon entropy within each bin, and the calibration error becomes the KL–calibration (Luo et al., 2025),*

$$R_T(p_{1:T}, y_{1:T}) = \sum_p n_T(p)H(\rho_T^p) \, , \quad K_T(p_{1:T}, y_{1:T}) = \sum_p n_T(p)\, \mathrm{KL}(\rho_T^p \| p) \, .$$

Motivated by this decomposition, Foster and Hart (2023) compare the learner to the external forecaster's refinement score and define the notions of *calibeating* and *multi-calibeating*.

Definition 2 (Calibeating and Multi-Calibeating) *A learner is* $\alpha(T)$-multi-calibeating *w.r.t. loss* ℓ *if for any external forecasts* $\{q_{1:T}^{(n)}\}_{n=1}^N$ *and outcomes* $y_{1:T}$*, the learner's predictions* $p_{1:T}$ *satisfy*

$$L_T(p_{1:T}, y_{1:T}) \leq R_T(q_{1:T}^{(n)}, y_{1:T}) + \alpha(T) \, , \forall n \in [N] \, . \tag{1}$$

We call $\alpha(T)$ *the multi-calibeating rate. We say the learner is multi-calibeating if* $\alpha(T) = o(T)$*. When (1) holds in expectation over the learner's randomness, we call* $\alpha(T)$ *the expected multi-calibeating rate. When there is only* $N = 1$ *external forecast, we simply say calibeating.*

We also introduce another performance measure called calibration.

Definition 3 (Calibration) *A learner is $\beta(T)$-calibrated w.r.t. loss ℓ if for any outcome sequences $y_{1:T}$ (and any external forecasts), the learner's predictions $p_{1:T}$ satisfy*

$$K_T(p_{1:T}, y_{1:T}) \le \beta(T). \tag{2}$$

We call $\beta(T)$ the calibration rate and say the algorithm is calibrated if $\beta(T) = o(T)$. When (2) holds in expectation over the algorithm's randomness, we call $\beta(T)$ the expected calibration rate.

Note that calibeating and calibration are incomparable in general. Calibeating only guarantees that the learner's loss is no larger than the forecaster's loss minus the forecaster's calibration error, i.e., it competes with the *external forecaster*'s refinement score. The learner's calibration error might not vanish if it itself attains a low refinement term.

3. Calibeating $=$ No-Regret Learning

This section considers the calibeating problem, i.e., when there is only one external forecast every round. Let $Q \coloneqq \{q_t : t \in [T]\}$ denote the set of distinct external forecast values that appear over the horizon.[2] Foster and Hart (2023) study the Brier and log losses and give algorithms with calibeating rate of $O(|Q| \log T)$. We recover and extend their results by reductions to no-regret learning.

First, define the regret of predictions $p_{1:T}$ under outcomes $y_{1:T}$ to be

$$\mathrm{Reg}_T(p_{1:T}, y_{1:T}) \coloneqq \sum_{t=1}^{T} \ell(p_t, y_t) - \min_{p \in \Delta_K} \sum_{t=1}^{T} \ell(p, y_t) \ .$$

We say an algorithm has (expected-)regret of $\alpha(T)$ if $\mathrm{Reg}_T(p_{1:T}, y_{1:T}) \le \alpha(T)$ always holds (in expectation). The following theorem shows that calibeating reduces to no-regret learning.

Theorem 4 *For any proper loss ℓ and any online algorithm A with regret $\alpha(T)$, where α is a concave function, Algorithm 1 is $|Q|\alpha(T/|Q|)$-calibeating.*

Proof The reduction partitions the rounds $t \in [T]$ by the external forecast value q_t, and runs an independent copy of the no-regret learner A for each forecast value. Formally, for any external forecast q that appeared at least once, run a separate copy of A, denoted as A_q, on the subset of rounds $\mathcal{I}_q \coloneqq \{t : q_t = q\}$. For each subsequence, we have

$$\sum_{t: q_t = q} \ell(p_t, y_t) - \min_{p \in \Delta_K} \sum_{t: q_t = q} \ell(p, y_t) \le \alpha(n_T(q)) \ .$$

Summing up over all the subsequences and by Definition 1, we have

$$\begin{aligned} L_T(p_{1:T}, y_{1:T}) - R_T(q_{1:T}, y_{1:T}) &\le \sum_q \alpha(n_T(q)) \\ &\le |Q|\alpha\left(\frac{\sum_q n_T(q)}{|Q|}\right) \qquad \text{(Jensen's inequality)} \\ &= |Q|\alpha\left(\frac{T}{|Q|}\right) \ , \end{aligned}$$

Algorithm 1: Calibeating by Bin-Wise No-Regret

Input: Online learner A.
for $t = 1$ **to** T **do**
    // prediction
    Observe external forecasts $q_t \in \Delta_K$.
    if A_{q_t} *is uninitialized* **then**
        Initialize a fresh copy $\mathsf{A}_{q_t} \leftarrow \mathsf{A}$.
    end
    Query A_{q_t} and obtain prediction $p_t \in \Delta_K$.
    // update
    Observe outcome y_t and incur loss $\ell(p_t, y_t)$.
    Update A_{q_t} with y_t.
end

which finishes the proof. ■

We note that common regret bounds obtained for standard online algorithms are all concave in T, e.g., they are often of the form $O((\log T)^\alpha T^\beta)$ for some $\alpha > 0$ and $\beta \in [0, 1)$. The algorithm of Foster and Hart (2023) can be recovered as a special case of Theorem 4, with the online algorithm being follow–the–leader (FTL) (with a standard interior restriction for the log loss to avoid the unbounded boundary). Moreover, Theorem 4 readily obtains $O(\sqrt{|Q|KT})$ for bounded proper losses (Luo et al., 2024) and $O(|Q| \log T)$ for exp-concave losses (Hazan et al., 2007), which encompasses Brier and log losses as special cases.

Corollary 5 *For bounded proper loss ℓ, instantiating the online learner in Theorem 4 with the follow-the-perturbed-leader algorithm of Luo et al. (2024) yields an algorithm with expected calibeating rate of $O(\sqrt{|Q|KT})$.*

We note that, inherited from the no-regret guarantees in Luo et al. (2024), the algorithm in Corollary 5 can actually achieve $O(\sqrt{|Q|KT})$ *simultaneously* for all bounded proper losses.

Definition 6 *A convex function ℓ is η-exp-concave if $\exp(-\eta\ell)$ is concave.*

Corollary 7 *For an η-exp-concave loss ℓ (e.g., Brier and log losses), instantiating the online learner in Theorem 4 with exponentially weighted online optimization (EWOO) (Hazan, 2016) yields an algorithm with an expected calibeating rate of $O(\frac{|Q|K}{\eta} \log(1 + \frac{T}{|Q|}))$. For the Brier loss, instantiating the online learner with follow-the-leader yields an algorithm with an expected calibeating rate of $O(|Q| \log(1 + \frac{T}{|Q|}))$.*

Besides the upper bound, we also prove that any lower bound for no-regret learning with a proper loss implies a lower bound for calibeating. Combining with Theorem 4, our results show that calibeating is minimax-equivalent to regret minimization. We defer the proof to Section A.1.

2. We also use Q to denote the set of possible external forecast values for lower bound results.

Theorem 8 *For any proper loss ℓ, denote the optimal regret bound as*

$$\beta(T) \coloneqq \inf_{\mathsf{A}} \sup_{y_{1:T}\in\mathcal{E}^T} \mathbb{E}_{p_{1:T}\sim\mathsf{A}}\left[\sum_{t=1}^{T}\ell\left(p_t, y_t\right) - \min_{p\in\Delta_K}\sum_{t=1}^{T}\ell\left(p, y_t\right)\right], \tag{3}$$

where A ranges over (possibly randomized) online algorithms. Then, every algorithm is at best $|Q|\beta(\lfloor T/|Q|\rfloor)$-calibeating.

Combining Theorem 8 with known regret lower bounds for bounded proper losses (Luo et al., 2024) and Brier and log losses (Cesa-Bianchi and Lugosi, 2006) yields the following.

Corollary 9 *There exist bounded proper losses with calibeating rate at least $\Omega(\sqrt{|Q|KT})$.*

Corollary 10 *For the Brier loss, the calibeating rate is at least $\Omega(|Q|\log(T/|Q|))$. For the log loss, the calibeating rate is at least $\Omega(|Q|K\log(T/|Q|))$.*

4. Multi-Calibeating = Calibeating + Expert Problem

Next, we consider the multi-calibeating problem. Foster and Hart (2023) obtain multi-calibeating rates of $O((N+|Q|)\log T)$ and $O(\sqrt{NT}+|Q|\log T)$, via Blackwell approachability and online linear regression. Lee et al. (2022) achieve a rate logarithmic in N, but polynomial in T (more precisely, $\tilde{O}(\sqrt{|Q|}(\log N)^{\frac{1}{4}}T^{\frac{3}{4}})$ for the optimal choice of parameters).

This section presents a simple reduction from multi-calibeating to the expert problem. Via that reduction, we achieve the optimal multi-calibeating rates.

Expert Problem. The interaction protocol in this problem is the same as in multi-calibeating: at each round t, the learner observes N expert predictions $\{p_t^{(n)}\}_n \subseteq \Delta_K$, and makes its own prediction $p_t \in \Delta_K$. An experts algorithm E achieves regret $\gamma(T)$ if for every sequence $\{(p_t^{(1:N)}, y_t)\}_{t=1}^T$,

$$\mathbb{E}\left[\sum_{t=1}^{T}\ell(p_t, y_t)\right] \le \min_{n\in[N]}\sum_{t=1}^{T}\ell(p_t^{(n)}, y_t) + \gamma(T), \tag{4}$$

where the expectation is over the randomness of E.

Comparing this definition of regret with the definition of multi-calibeating rate, the only difference is that the latter remaps the experts/forecasters' predictions optimally, while the former does not. The remapping of each individual forecaster is precisely the problem of calibeating. Hence, we run a separate calibeating algorithm for each forecaster, and use an experts algorithm to aggregate their decision. See Algorithm 2 for a formal description.

Theorem 11 *For any loss function ℓ, any calibeating algorithm with rate $\alpha(T)$, and any experts algorithm with regret $\gamma(T)$, Algorithm 2 is $(\alpha(T)+\gamma(T))$-calibeating.*

Proof By the regret bound of algorithm E, for any $n \in [N]$, we have

$$\mathbb{E}\left[L_T(p_{1:T}, y_{1:T})\right] \le L_T(p_{1:T}^{(n)}, y_{1:T}) + \gamma(T)\,.$$

Algorithm 2: Multicalibeating by Expert Aggregation

Sub-routines:

- For each forecaster $n \in [N]$, a separate calibeating algorithm $\mathsf{A}^{(n)}$ (Algorithm 1).
- Experts algorithm E (e.g., Hedge, Freund and Schapire, 1997).

for $t = 1$ **to** T **do**

  `// prediction`

  Observe external forecasts $q_t^{(1)}, \dots, q_t^{(N)} \in \Delta_K$.

  For each $n \in [N]$, query $\mathsf{A}^{(n)}$ with forecast $q_t^{(n)}$ to get its prediction $p_t^{(n)} \in \Delta_K$.

  Query E with $\{p_t^{(n)}\}_{n=1}^N$ as the experts' forecasts, and follow its prediction p_t.

  `// update`

  Observe outcome y_t and update $\mathsf{A}^{(n)}$ for each $n \in [N]$ with this outcome.

  Update E with $\ell(p_t^{(n)}, y_t)$ as the loss of expert $n \in [N]$.

end

By the calibeating rate of algorithm A, we have

$$L_T(p_{1:T}^{(n)}, y_{1:T}) \le R_T(q_{1:T}^{(n)}, y_{1:T}) + \alpha(T).$$

Combining these inequalities yields a multi-calibeating rate of $\alpha(T) + \gamma(T)$. ■

Let $Q^{(n)} := \{q_t^{(n)} : t \in [T]\}$ denote the set of distinct external forecasts made by forecaster $n \in [N]$. We assume $|Q^{(n)}| = |Q|$ for all n for simplicity.[3] By the regret bounds of Hedge (e.g., Bubeck, 2011, Theorems 2.1 and 2.2), and the calibeating rates in Corollaries 5 and 7, we get the following corollaries. In contrast to the loss-oblivious property of Corollary 5, the algorithm in Corollary 12 requires a fixed ℓ, as the experts algorithm relies on the loss values to update.

Corollary 12 *For any bounded proper loss ℓ, there exists an algorithm with an expected multi-calibeating rate of $O(\sqrt{T \log N} + \sqrt{|Q|KT})$.*

Corollary 13 *For any η-exp-concave loss ℓ (e.g., Brier and Log losses), there exists an algorithm with an expected multi-calibeating rate of $O(\frac{1}{\eta}(\log N + |Q|K \log(1 + \frac{T}{|Q|})))$. For the Brier loss, there exists an algorithm with an expected rate of $O(\log N + |Q| \log(1 + \frac{T}{|Q|}))$.*

We show a minimax-equivalence of the two problems. Since now external forecasts are involved, we consider calibeating rates and expert regrets as functions of both the time round and the number of possible distinct external forecast values. Similarly to $Q^{(n)}$, given an instance of the expert problem, let $P^{(n)} := \{p_t^{(n)} : t \in [T]\}$ denote the set of possible expert predictions made by expert forecaster $n \in [N]$. We have the following theorem.

Theorem 14 *For any proper loss ℓ, suppose there exist functions $\phi, \lambda : \mathbb{Z}^2 \to \mathbb{R}$ such that for any T and m,*

$$\inf_{\mathsf{A}} \sup_{\substack{(q_{1:T}, y_{1:T}): \\ |Q| \le m}} \mathbb{E}_{p_{1:T} \sim \mathsf{A}} \left[L_T(p_{1:T}, y_{1:T}) - R_T(q_{1:T}, y_{1:T}) \right] \ge \phi(T, m),$$

3. Our results also hold when $|Q^{(n)}|$s differ across forecasters, and the resulting bounds adapt to specific forecasters.

where A *ranges over all randomized calibeating algorithms, and,*

$$\inf_{\mathsf{E}} \sup_{\substack{(p_{1:T}^{(1:N)}, y_{1:T}): \\ \forall n, |P^{(n)}| \leq m}} \mathbb{E}_{p_{1:T} \sim \mathsf{E}} \left[\sum_{t=1}^{T} \ell(p_t, y_t) - \min_{n \in [N]} \sum_{t=1}^{T} \ell(p_t^{(n)}, y_t) \right] \geq \lambda(T, m),$$

where E *ranges over all randomized expert algorithms. Then,*

$$\inf_{\mathsf{M}} \sup_{\substack{(q_{1:T}^{(1:N)}, y_{1:T}): \\ \forall n, |Q^{(n)}| \leq m}} \mathbb{E}_{p_{1:T} \sim \mathsf{M}} \left[L_T(p_{1:T}, y_{1:T}) - \min_{n \in [N]} R_T(q_{1:T}^{(n)}, y_{1:T}) \right] \geq \max \left\{ \phi(T, m), \lambda(T, m) \right\},$$

where M *ranges over all multi-calibeating algorithms.*

By known lower bounds for expert problems when the expert predictions can be arbitrary values (Cesa-Bianchi and Lugosi, 2006), and that the lower-bound examples can be obtained when the size of distinct expert prediction values is constant, we show the following lower bounds for multi-calibeating, matching the upper bounds.

Corollary 15 *There exist bounded proper losses under which the multi-calibeating rates are at least* $\Omega(\sqrt{T \log N} + \sqrt{|Q| K T})$.

Corollary 16 *For the Brier loss, the multi-calibeating rate of any algorithm is at least* $\Omega(\log N + |Q| \log(T/|Q|))$. *For the log loss, the multi-calibeating rate is at least* $\Omega(\log N + |Q| K \log(T/|Q|))$.

5. Calibeating and Calibration at the Same Time

In this section, we consider the problem of achieving simultaneous calibeating and calibration. Existing approaches focus on the Brier loss. Foster and Hart (2023) obtain simultaneous rates of $\tilde{O}(|Q|^{\frac{2}{K+1}} T^{\frac{K-1}{K+1}})$ via bin refinement and stochastic fixed-point methods, while Lee et al. (2022) obtain $\tilde{O}(\sqrt{|Q|} (\log N)^{\frac{1}{4}} T^{\frac{3}{4}})$ in the binary case after parameter tuning.

We focus on the Brier loss and provide new and improved bounds for simultaneous calibeating and calibration. Specifically, we provide a meta-algorithm (Algorithm 3) which, for any given external reference algorithm A^*, keeps careful track of the losses of A^* while ensuring calibration.

Theorem 17 *For Brier loss, any* $\varepsilon \in (0, 1)$, *and any reference algorithm* A^*, *Algorithm 3 simultaneously guarantees an expected regret of at most* $O(\varepsilon^2 T)$ *compared to* A^*, *and a calibration rate of at most* $O_{K, \log T}(\sqrt{T} + \frac{1}{\varepsilon^{K-1}} \log \frac{1}{\varepsilon} + \varepsilon^2 T)$ *with high probability.*

Here, the $O_{K, \log T}$ notation hides a factor polynomial in K and $\log T$ for readability. We hide the $\log T$ factors because for the calibration error, the dominant dependence in T is polynomial, and we hide the K factors because the bounds degenerate to the trivial $O(T)$ when K gets larger and larger. The bounds from previous works are also polynomial in K.

Let A^* be a multi-calibeating algorithm from the previous sections. With $\varepsilon = \sqrt{\frac{\log T}{T}}$, for $K = 2$, we obtain the optimal calibeating rate, improving the polynomial-in-T dependence in Foster and Hart (2023).

Corollary 18 *For Brier loss with binary outcomes, there is an algorithm with an expected multi-calibeating rate of at most* $O(\log N + |Q| \log T)$*, and a calibration rate of at most* $O_{\log T}(\sqrt{T})$ *with high probability.*

With $\varepsilon = (\frac{\log T}{T})^{\frac{1}{K+1}}$ for $K \geq 3$, we achieve the same calibration rate as in Fishelson et al. (2025); Foster and Hart (2023), and drop the $|Q|-$dependence in Foster and Hart (2023).

Corollary 19 *For Brier loss and* $K \geq 3$ *outcomes, there is an algorithm with an expected multi-calibeating rate of at most* $O_{K,\log T}(\log N + |Q| + T^{\frac{K-1}{K+1}})$*, and a calibration rate of at most* $O_{K,\log T}(T^{\frac{K-1}{K+1}})$ *with high probability.*

For $K \geq 3$ outcomes, we can also lower the multi-calibeating rate at the cost of raising the calibration rate, by choosing a different ε.

Corollary 20 *For Brier loss and* $K \geq 3$ *outcomes, for any* $x \in (\frac{K-3}{K-1}, \frac{K-1}{K+1}]$*, there is an algorithm with expected multi-calibeating rate of at most* $O_{K,\log T}(\log N + |Q| + T^x)$*, and a calibration rate of at most* $O_{K,\log T}(T^{\frac{(K-1)(1-x)}{2}})$ *with high probability.*

5.1. Algorithm

Discretization and rounding. To achieve calibration, it is necessary to focus on a finite set of predictions via discretization. For that, we consider a triangulation of Δ_K and randomly round each prediction to a vertex of the triangulation (recall that ℓ is fixed to the Brier loss in this section).

Lemma 21 (Fishelson et al. 2025) *For any* $\varepsilon \in (0, 1)$*, there is a subset of predictions* $\mathcal{K}^\varepsilon \subset \Delta_K$ *of size* $M = |\mathcal{K}^\varepsilon| = O(\sqrt{K}\,\varepsilon^{-K+1})$*, and a rounding scheme* $\mathsf{H} : \Delta_K \to \Delta(\mathcal{K}^\varepsilon)$ *that maps an arbitrary prediction* $q \in \Delta_K$ *to a distribution over those in* $\mathcal{K}^\varepsilon$*, such that for any outcome* $y \in \mathcal{E}$*, we have* $\mathbb{E}_{s\sim\mathsf{H}(q)}[\ell(s, y)] \leq \ell(q, y) + O(\varepsilon^2)$*.*

Blum-Mansour reduction. For the connection between calibration and no-swap-regret learning, we employ the well-known reduction by Blum and Mansour (2007) with the $O(\log T)$ regret online learning algorithm for Brier loss, e.g., FTL. We present the algorithm and its proof in Section C.2.

Lemma 22 *There is an online algorithm* A_{BM} *that, in each step* $t \in [T]$*, first predicts an* $M \times M$ *column-stochastic matrix* A_t*, and then observes outcome* y_t *and a distribution* $\pi_t \in \Delta(\mathcal{K}^\varepsilon)$*, such that for any transformation* $\sigma : \Delta_K \to \Delta_K$*,*

$$\sum_{t\in[T]} \mathbb{E}_{p_t\sim A_t\pi_t}\ell(p_t, y_t) \leq \sum_{t\in[T]} \mathbb{E}_{p'_t\sim\pi_t}\ell(\sigma(p'_t), y_t) + O\left(M \log T + \varepsilon^2 T\right) .$$

Intuitively, we may interpret the column-stochastic matrix A_t from algorithm A_{BM} as a suggested remapping from any prediction, so that for any sequence of outcomes y_t and distributions of predictions π_t, the remapped/calibrated predictions are competitive against the best remapping σ in hindsight. The standard approach is then to sample a randomized prediction from the stationary distribution of A_t (but we will do this step later after mixing A_t with another remapping matrix).

Interpolating between calibration and multi-calibeating. Besides achieving small swap regret and calibration rate, we also want to follow the reference prediction b_t from algorithm A^* to be competitive against this reference algorithm. Observe that following the reference prediction corresponds to remapping every prediction to b_t, which can be captured by a remapping matrix $B_t = (b_t, b_t, \ldots, b_t) \in \mathbb{R}^{M \times M}$. To hedge between these two factors, we resort to a lopsided two-expert algorithm $\mathsf{A}_{\text{lopsided}}$ to obtain a weight $w_t \in [0, 1]$, and take a linear combination $C_t = w_t A_t + (1 - w_t) B_t$ as the aggregated remapping.

Lemma 23 (Sani et al. 2014) *There is an algorithm $\mathsf{A}_{\text{lopsided}}$ for the expert problem with two experts, such that the expected regret w.r.t. expert 1 is at most $O(\sqrt{T \log T})$, and the expected regret w.r.t. expert 2 is at most $O(1)$.*

Finally, we sample a prediction from the stationary distribution of C_t, as shown in Algorithm 3.

Algorithm 3: Multi-Calibeating + Calibration

Sub-routines:

- Discretization and rounding algorithm H (see Fishelson et al., 2025)
- Reference algorithm A^* (Algorithm 1 for calibeating, Algorithm 2 for multi-calibeating).
- BM reduction A_{BM} (Algorithm 3.1).
- Lopsided two-expert algorithm $\mathsf{A}_{\text{lopsided}}$ (Algorithm 3.2).

for $t = 1$ **to** T **do**

```
// prediction
```

Algorithm A_{BM} predicts A_t.
Round algorithm A^*'s prediction with H to get $b_t \in \Delta(\mathcal{K}^\varepsilon)$, and let $B_t = (b_t, \ldots, b_t)$.
Algorithm $\mathsf{A}_{\text{lopsided}}$ predicts $w_t \in [0, 1]$.
Let $C_t = w_t A_t + (1 - w_t) B_t$, and $\pi_t \in \Delta_M$ be its stationary distribution, i.e., $\pi_t = C_t \pi_t$.
Predict $p_t \sim \pi_t$.

```
// update
```

Observe outcome y_t.
Update A_{BM} and A^* based on y_t and π_t (applicable to the former).
Update $\mathsf{A}_{\text{lopsided}}$ with $\mathbb{E}_{z \sim A_t \pi_t} \ell(z, y_t)$ and $\mathbb{E}_{z \sim b_t} \ell(z, y_t)$ as the losses of experts 1 and 2.

end

5.2. Analysis: Proof of a weaker version of Theorem 17

We will prove a weaker guarantee of pseudo-calibration due to space constraints and defer the rest of the proof to Section C.4. By definition, the expected cumulative loss of Algorithm 3 is

$$\mathbb{E}\, L_T(p_{1:T}, y_{1:T}) = \mathbb{E}_{\pi_{1:T}} \left[\sum_{t \in [T]} \mathbb{E}_{z_t \sim \pi_t} \ell(z_t, y_t) \right].$$

Since π_t is the stationary distribution of $C_t = w_t A_t + (1-w_t)B_t$, the above further equals

$$\sum_{t\in[T]} \mathbb{E}_{z_t\sim C_t\pi_t}\ell(z_t,y_t) = \sum_{t\in[T]} \left(w_t\,\mathbb{E}_{z_t\sim A_t\pi_t}\ell(z_t,y_t) + (1-w_t)\,\mathbb{E}_{z_t\sim B_t\pi_t}\ell(z_t,y_t)\right)$$
$$= \sum_{t\in[T]} \left(w_t\,\mathbb{E}_{z_t\sim A_t\pi_t}\ell(z_t,y_t) + (1-w_t)\,\mathbb{E}_{z_t\sim b_t}\ell(z_t,y_t)\right) .$$

By construction, $\mathbb{E}_{z_t\sim A_t\pi_t}\ell(z_t,y_t)$ and $\mathbb{E}_{z_t\sim b_t}\ell(z_t,y_t)$ are the losses of the two-expert problem in round $t\in[T]$. The lopsided regret bounds of $\mathsf{A}_{\text{lopsided}}$ (Lemma 23) give

$$\mathbb{E}\,L_T(p_{1:T},y_{1:T}) \le \mathbb{E}_{b_{1:T}}\left[\sum_{t\in[T]} \mathbb{E}_{z_t\sim A_t\pi_t}\ell(z_t,y_t)\right] + O(\sqrt{T\log T})\,, \tag{5}$$

$$\mathbb{E}\,L_T(p_{1:T},y_{1:T}) \le \mathbb{E}_{b_{1:T}}\left[\sum_{t\in[T]} \mathbb{E}_{z_t\sim b_t}\ell(z_t,y_t)\right] + O(1)\,. \tag{6}$$

Regret w.r.t. A^*. Recall that b_t is obtained by rounding the prediction from A^* with rounding algorithm H. By the $O(\varepsilon^2)$ rounding error bound of H (Lemma 21) and Eq. (6), the regret w.r.t. the reference algorithm A^* is at most $O(\varepsilon^2 T)$.

Pseudo-calibration. We will prove a weaker guarantee that

$$\mathbb{E}\,L_T(p_{1:T},y_{1:T}) \le \min_{\sigma:\Delta_K\to\Delta_K} \mathbb{E}\sum_{t=1}^{T} \ell(\sigma(p_t),y_t) + O\left(\sqrt{T\log T} + \frac{\sqrt{K}}{\varepsilon^{K-1}}\log T + \varepsilon^2 T\right) .$$

This follows from Eq. (5), the guarantee of BM reduction (Lemma 22), and that $M = O(\sqrt{K}\varepsilon^{-K+1})$ (Lemma 21). It is weaker than the original statement as the choice of σ does not depend on the realization of randomness of the algorithm. By contrast, the original statement allows choosing σ based on the realization of randomness. We defer this concentration argument to Section C.4.

6. Conclusion and Discussion

We have revisited calibeating through the lens of online learning and developed a reduction-based framework that connects calibeating and its extensions to standard online-learning notions. This viewpoint enables us to recover and sharpen existing results, extend them to general proper scoring losses, and deliver new matching lower bounds, in a unified and modular way.

Dependence on $|Q|$. A recurring feature of our bounds is their dependence on $|Q|$, the number of distinct external forecast values. Our matching upper and lower bounds suggest that this dependence is intrinsic to the classical formulation of calibeating rather than an artifact of the analysis. Intuitively, the refinement benchmark decomposes over distinct forecast values, so a larger $|Q|$ induces a finer partition of the time horizon and therefore a stronger benchmark. In the extreme case of $|Q| = T$, each forecast value forms a singleton bin, making the refinement benchmark particularly demanding. This regime is natural in decision-oriented settings, where forecasts are often discrete or quantized into a relatively small menu of values. For example, weather forecasts in phone apps are often displayed on a coarse grid (Fishelson et al., 2025). Therefore, to reduce the dependence on $|Q|$ implies the need of a different notion, and studying discretized or approximate variants of calibeating for highly continuous forecast spaces is an interesting direction for future work.

Beyond the Brier loss. While our simultaneous calibration and calibeating results are stated for the Brier loss, the meta-algorithm developed in Section 5 is more modular than it may initially appear. At a high level, the framework relies on two ingredients: a discretization and rounding scheme with controlled loss guarantees, and a mechanism for upgrading pseudo-swap-regret guarantees to true swap-regret guarantees. For the Brier loss, both ingredients are available through existing results. Extending the framework to other proper losses therefore reduces largely to establishing analogous ingredients. In particular, it would be interesting to determine whether similar simultaneous guarantees can be obtained for broader classes of proper losses, such as bounded proper losses and the log loss.

Other open problems. A natural direction for future work is to push this approach further, both to identify additional achievable guarantees and to improve online forecasting more broadly. Another natural open question is whether one can simultaneously achieve the optimal $O(|Q|\log T)$ calibeating rate and the best known $\tilde{O}(T^{\frac{1}{3}})$ calibration rate in the binary case.

Acknowledgments

We thank Nishant Mehta for pointing out the missing K-dependence in our bounds for the log loss. Haipeng Luo is supported by NSF award IIS-1943607. Zhiyi Huang is supported in part by RGC Hong Kong (No. 17211723). This publication is part of the Chair "Markets and Learning", supported by Air Liquide, BNP PARIBAS ASSET MANAGEMENT Europe, EDF, Orange and SNCF, sponsors of the Inria Foundation. This work has also received support from the French government, managed by the National Research Agency, under the France 2030 program with the reference "PR[AI]RIE-PSAI" (ANR-23-IACL-0008). Yurong Chen is supported by a Marie Skłodowska-Curie Fellowship from the European Union's Horizon Europe Research and Innovation program (Grant Agreement No. 101272753).

References

Alina Beygelzimer, John Langford, Lihong Li, Lev Reyzin, and Robert Schapire. Contextual bandit algorithms with supervised learning guarantees. In Geoffrey Gordon, David Dunson, and Miroslav Dudík, editors, *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, volume 15 of *Proceedings of Machine Learning Research*, pages 19–26, Fort Lauderdale, FL, USA, 11–13 Apr 2011. PMLR. URL https://proceedings.mlr.press/v15/beygelzimer11a.html.

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

Jochen Bröcker. Reliability, sufficiency, and the decomposition of proper scores. *Quarterly Journal of the Royal Meteorological Society*, 135(643):1512–1519, 2009. doi: https://doi.org/10.1002/qj.456. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.456.

Sébastien Bubeck. Introduction to online optimization, 12 2011. URL https://www.microsoft.com/en-us/research/publication/introduction-online-optimization/.

Nicolo Cesa-Bianchi and Gabor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

Philip Dawid. Probability forecasting. In *Encyclopedia of Statistical Sciences*. Wiley and Sons, New York, NY, 2006.

Shachi Deshpande, Charles Marx, and Volodymyr Kuleshov. Calibrated regression against an adversary without regret, 2024. URL https://arxiv.org/abs/2302.12196.

Maxwell Fishelson, Robert Kleinberg, Princewill Okoroafor, Renato Paes Leme, Jon Schneider, and Yifeng Teng. Full swap regret and discretized calibration. In Gautam Kamath and Po-Ling Loh, editors, *Proceedings of The 36th International Conference on Algorithmic Learning Theory*, volume 272 of *Proceedings of Machine Learning Research*, pages 444–480. PMLR, 2 2025. URL https://proceedings.mlr.press/v272/fishelson25a.html.

Dean Foster and Sergiu Hart. "calibeating": beating forecasters at their own game. *Theoretical Economics*, 18(4):1441–1474, 2023. URL https://econtheory.org/ojs/index.php/te/article/view/20231441/0.

Dean P. Foster and Rakesh V. Vohra. Asymptotic calibration. *Biometrika*, 85(2):379–390, 1998. ISSN 0006-3444,1464-3510. doi: 10.1093/biomet/85.2.379. URL https://doi.org/10.1093/biomet/85.2.379.

Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.

Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberger. On calibration of modern neural networks. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 1321–1330. PMLR, 8 2017. URL https://proceedings.mlr.press/v70/guo17a.html.

Chirag Gupta and Aaditya Ramdas. Online platt scaling with calibeating. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 12182–12204. PMLR, 7 2023. URL https://proceedings.mlr.press/v202/gupta23c.html.

Elad Hazan. Introduction to online convex optimization. *Foundations and Trends in Optimization*, 2(3-4):157–325, 08 2016. ISSN 2167-3888. doi: 10.1561/2400000013. URL https://doi.org/10.1561/2400000013.

Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2):169–192, 2007.

Volodymyr Kuleshov and Stefano Ermon. Estimating uncertainty online against an adversary. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

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

Haipeng Luo, Spandan Senapati, and Vatsal Sharan. Optimal multiclass u-calibration error and beyond. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, editors, *Advances in Neural Information Processing Systems*, volume 37, pages 7521–7551. Curran Associates, Inc., 2024. doi: 10.52202/079017-0241. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/0e4d695de4c606494ba9b0f3dac3b57a-Paper-Conference.pdf.

Haipeng Luo, Spandan Senapati, and Vatsal Sharan. Simultaneous swap regret minimization via kl-calibration, 2025. URL https://arxiv.org/abs/2502.16387.

Charles Marx, Volodymyr Kuleshov, and Stefano Ermon. Calibrated probabilistic forecasts for arbitrary sequences, 2025. URL https://arxiv.org/abs/2409.19157.

Princewill Okoroafor, Bobby Kleinberg, and Wen Sun. Faster recalibration of an online predictor via approachability. In Sanjoy Dasgupta, Stephan Mandt, and Yingzhen Li, editors, *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pages 4690–4698. PMLR, 5 2024. URL https://proceedings.mlr.press/v238/okoroafor24a.html.

Frederick Sanders. On subjective probability forecasting. *Journal of Applied Meteorology and Climatology*, 2(2):191 – 201, 1963. doi: 10.1175/1520-0450(1963)002⟨0191:OSPF⟩2.0.CO;2. URL https://journals.ametsoc.org/view/journals/apme/2/2/1520-0450_1963_002_0191_ospf_2_0_co_2.xml.

Amir Sani, Gergely Neu, and Alessandro Lazaric. Exploiting easy data in online optimization. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014. URL https://proceedings.neurips.cc/paper_files/paper/2014/file/8cfb39d9174128beb141866808bd154e-Paper.pdf.

Appendix A. Omitted Proofs in Section 3

A.1. Proof of Theorem 8

Theorem 8 *For any proper loss ℓ, denote the optimal regret bound as*

$$\beta(T) := \inf_{\mathsf{A}} \sup_{y_{1:T} \in \mathcal{E}^T} \mathbb{E}_{p_{1:T} \sim \mathsf{A}} \left[\sum_{t=1}^{T} \ell\left(p_t, y_t\right) - \min_{p \in \Delta_K} \sum_{t=1}^{T} \ell\left(p, y_t\right) \right], \tag{3}$$

where A *ranges over (possibly randomized) online algorithms. Then, every algorithm is at best* $|Q|\beta(\lfloor T/|Q|\rfloor)$*-calibeating.*

Proof We prove a general result that, for any integers $T_q \geq T_0$ with $\sum_{q\in Q} T_q = T$, we have,

$$\inf_{\mathsf{A}} \sup_{(q_{1:T}, y_{1:T})} \mathbb{E}_{p_{1:T}\sim\mathsf{A}}\left[L_T(p_{1:T}, y_{1:T}) - R_T(q_{1:T}, y_{1:T})\right] \geq \sum_{q\in Q} \beta\left(T_q\right).$$

By Yao's minimax principle, (3) implies that, for any T, there exists a distribution $\mathrm{S}_T \in \Delta(\mathcal{E}^T)$ such that

$$\min_{\mathsf{D}} \mathbb{E}_{y_{1:T}\sim \mathrm{S}_T}\left[\sum_{t=1}^{T} \ell\left(p_t, y_t\right) - \min_{p\in\Delta_K} \sum_{t=1}^{T} \ell\left(p, y_t\right)\right] \geq \beta(T), \tag{7}$$

where the minimum is over deterministic online algorithms D.

For any q and the corresponding T_q, let $\mathrm{S}_q = \mathrm{S}_{T_q}$ be the distribution guaranteed by (7) at horizon T_q. Define the following distribution S over pairs$(q_{1:T}, y_{1:T}) \in (Q \times \mathcal{E})^T$ as follows:

1. Choose disjoint index sets $\{\mathcal{I}_q\}_{q\in Q}$ with $\mathcal{I}_q \subseteq [T]$, $|\mathcal{I}_q| = T_q$ and $\cup_{q\in Q}\mathcal{I}_q = [T]$. Set $q_t = q$ for all $t \in \mathcal{I}_q$.
2. For each $q \in Q$, denote the subsequence of outcomes y_ts that $t \in \mathcal{I}_q$ to be $y_{\mathcal{I}_q} = (y_t)_{t\in\mathcal{I}_q}$. Independently sample $y_{\mathcal{I}_q}$ according to S_q.

Then by definition and the additivity of sums of expectations,

$$\begin{aligned}
&\min_{\mathsf{D}} \mathbb{E}_{(q_{1:T}, y_{1:T})\sim \mathrm{S}}\left[L_T(p_{1:T}, y_{1:T}) - R_T(q_{1:T}, y_{1:T})\right] \\
&\quad = \min_{\mathsf{D}} \mathbb{E}_{(q_{1:T}, y_{1:T})\sim \mathrm{S}}\left[\sum_{q\in Q}\left(\sum_{t:q_t=q} \ell(p_t, y_t) - \min_{p\in\Delta_K}\sum_{t:q_t=q} \ell(p, y_t)\right)\right] \\
&\quad = \min_{\mathsf{D}} \sum_{q\in Q} \mathbb{E}_{y_{\mathcal{I}_q}\sim S_q}\left[\sum_{t:q_t=q} \ell(p_t, y_t) - \min_{p\in\Delta_K}\sum_{t:q_t=q} \ell(p, y_t)\right] \\
&\quad \geq \sum_{q\in Q} \min_{\mathsf{D}} \mathbb{E}_{y_{\mathcal{I}_q}\sim S_q}\left[\sum_{t:q_t=q} \ell(p_t, y_t) - \min_{p\in\Delta_K}\sum_{t:q_t=q} \ell(p, y_t)\right] \\
&\quad \geq \sum_{q\in Q} \beta(T_q).
\end{aligned}$$

Therefore, since a randomized algorithm A is a probability distribution over deterministic algorithms, it holds that

$$\begin{aligned}
&\min_{\mathsf{A}} \sup_{(q_{1:T}, y_{1:T})} \mathbb{E}_{p_{1:T}\sim\mathsf{A}}\left[L_T(p_{1:T}, y_{1:T}) - R_T(q_{1:T}, y_{1:T})\right] \\
&\quad \geq \min_{\mathsf{A}} \mathbb{E}_{(q_{1:T}, y_{1:T})\sim \mathrm{S}}\left[L_T(p_{1:T}, y_{1:T}) - R_T(q_{1:T}, y_{1:T})\right] \geq \sum_{q\in Q} \beta(T_q).
\end{aligned}$$

Choosing balanced $T_q \in \{\lfloor T/|Q|\rfloor, \lceil T/|Q|\rceil\}$ completes the proof. ■

Appendix B. Omitted Proofs in Section 4

B.1. Proof of Theorem 14

Theorem 14 *For any proper loss ℓ, suppose there exist functions $\phi, \lambda : \mathbb{Z}^2 \to \mathbb{R}$ such that for any T and m,*

$$\inf_{\mathsf{A}} \sup_{\substack{(q_{1:T}, y_{1:T}): \\ |Q| \le m}} \mathbb{E}_{p_{1:T} \sim \mathsf{A}} \left[L_T(p_{1:T}, y_{1:T}) - R_T(q_{1:T}, y_{1:T}) \right] \ge \phi(T, m),$$

where A ranges over all randomized calibeating algorithms, and,

$$\inf_{\mathsf{E}} \sup_{\substack{(p_{1:T}^{(1:N)}, y_{1:T}): \\ \forall n, |P^{(n)}| \le m}} \mathbb{E}_{p_{1:T} \sim \mathsf{E}} \left[\sum_{t=1}^{T} \ell(p_t, y_t) - \min_{n \in [N]} \sum_{t=1}^{T} \ell(p_t^{(n)}, y_t) \right] \ge \lambda(T, m),$$

where E ranges over all randomized expert algorithms. Then,

$$\inf_{\mathsf{M}} \sup_{\substack{(q_{1:T}^{(1:N)}, y_{1:T}): \\ \forall n, |Q^{(n)}| \le m}} \mathbb{E}_{p_{1:T} \sim \mathsf{M}} \left[L_T(p_{1:T}, y_{1:T}) - \min_{n \in [N]} R_T(q_{1:T}^{(n)}, y_{1:T}) \right] \ge \max \left\{ \phi(T, m), \lambda(T, m) \right\},$$

where M ranges over all multi-calibeating algorithms.

Proof First, for any realization of expert predictions $\{p_t^{(n)}\}_{t \in [T], n \in [N]}$ and outcomes $y_{1:T}$, consider the multi-calibeating problem with the same outcome sequence and the external forecasts $q_t^{(n)} = p_t^{(n)}$ for any $t \in [T]$ and $n \in [N]$ (Recall that the experts problem and multi-calibeating only differ in the benchmarks). By Definition 1,

$$R_T(q_{1:T}^{(n)}, y_{1:T}) = \sum_{p} \min_{u \in \Delta_K} \sum_{t \le T:\, q_t^{(n)} = p} \ell(u, y_t) \le \sum_{p} \sum_{t \le T:\, q_t^{(n)} = p} \ell(p, y_t) = \sum_{t=1}^{T} \ell(q_t^{(n)}, y_t),$$

where the inequality chooses $u = p$ in each bin. Taking $\min_{n \in [N]}$ on both sides yields

$$\min_{n \in [N]} R_T(q_{1:T}^{(n)}, y_{1:T}) \le \min_{n \in [N]} \sum_{t=1}^{T} \ell(q_t^{(n)}, y_t).$$

Therefore, for any predictions $p_{1:T}$,

$$\begin{aligned} L_T(p_{1:T}, y_{1:T}) - \min_{n \in [N]} R_T(q_{1:T}^{(n)}, y_{1:T}) &\ge \sum_{t=1}^{T} \ell(p_t, y_t) - \min_{n \in [N]} \sum_{t=1}^{T} \ell(q_t^{(n)}, y_t) \\ &= \sum_{t=1}^{T} \ell(p_t, y_t) - \min_{n \in [N]} \sum_{t=1}^{T} \ell(p_t^{(n)}, y_t), \end{aligned}$$

and multi-calibeating inherits the lower bound of the expert problem.

Second, consider the instances where $q_t^{(n)} = q_t$ for all $t \in [T]$, which is equivalent to beating only one external forecaster. Therefore, multi-calibeating inherits the lower bound of the calibeating problem. This completes the proof.

■

Appendix C. Omitted Proofs in Section 5

C.1. Proof of Corollary 20

Corollary 20 *For Brier loss and $K \geq 3$ outcomes, for any $x \in (\frac{K-3}{K-1}, \frac{K-1}{K+1}]$, there is an algorithm with expected multi-calibeating rate of at most $O_{K,\log T}(\log N + |Q| + T^x)$, and a calibration rate of at most $O_{K,\log T}(T^{\frac{(K-1)(1-x)}{2}})$ with high probability.*

Proof Let $\varepsilon^2 T = T^x$, then the calibeating rate is $O(|Q^{(n)}| \log T + T^x)$ and $\varepsilon = T^{\frac{x-1}{2}}$. The corresponding calibration error is of the order $O(\sqrt{T \log T} + T^{\frac{(K-1)(1-x)}{2}} \log T + T^x)$, which always has greater order than the regret. As the order of calibration error is optimized when $x = \frac{(K-1)(1-x)}{2} = \frac{K-1}{K+1}$, we consider smaller x with higher $T^{\frac{(K-1)(1-x)}{2}}$ term. The tradeoff follows by solving $T^{\frac{(K-1)(1-x)}{2}} < T$. ■

C.2. Algorithm A_{BM} and the proof of Lemma 22

Subroutine 3.1: A_{BM}

Sub-routines:

- For each grid action z_i, a separate online learner $\mathsf{A}^{(i)}$ (e.g., FTL)
- Discretization and rounding algorithm H (see Fishelson et al., 2025)

for $t = 1$ **to** T **do**
    `// prediction`
    for $i = 1$ **to** M **do**
        Observe strategy $q_t^{(i)} \in \Delta_K$ from Algorithm $\mathsf{A}^{(i)}$.
        Round $q_t^{(i)}$ to $\mathsf{H}(q_t^{(i)}) \in \Delta(\mathcal{K}^\varepsilon)$.
    end
    Output $A_t = (\mathsf{H}(q_t^{(1)}), \ldots, \mathsf{H}(q_t^{(M)})) \in \mathbb{R}^{M \times M}$.
    `// update`
    Receive feedback tuple (y_t, π_t).
    Update Algorithm $\mathsf{A}^{(i)}$ with loss function $\pi_t(i)\ell(\cdot, y_t)$.
end

Lemma 22 *There is an online algorithm A_{BM} that, in each step $t \in [T]$, first predicts an $M \times M$ column-stochastic matrix A_t, and then observes outcome y_t and a distribution $\pi_t \in \Delta(\mathcal{K}^\varepsilon)$, such that for any transformation $\sigma : \Delta_K \to \Delta_K$,*

$$\sum_{t \in [T]} \mathbb{E}_{p_t \sim A_t \pi_t} \ell(p_t, y_t) \leq \sum_{t \in [T]} \mathbb{E}_{p'_t \sim \pi_t} \ell(\sigma(p'_t), y_t) + O\left(M \log T + \varepsilon^2 T\right) .$$

Proof Denote $\mathcal{K}^\varepsilon := \{z_1, \ldots, z_M\}$. For any $j \in [M]$ and $\sigma(z_j) \in \Delta_K$, we have

$$\begin{aligned}
&\sum_t \pi_t(j) \sum_i A_t(i,j)\ell(z_i, y_t) - \sum_t \pi_t(j)\ell(\sigma(z_j), y_t) \\
&\quad = \sum_t \pi_t(j)\mathbb{E}_{i\sim \mathsf{H}(q_t^{(j)})}\ell(z_i, y_t) - \sum_t \pi_t(j)\ell(\sigma(z_j), y_t) \\
&\quad \overset{(a)}{\le} \sum_t \pi_t(j)\left(\ell(q_t^{(j)}, y_t) + C_2\varepsilon^2\right) - \sum_t \pi_t(j)\ell(\sigma(z_j), y_t) \\
&\quad = \sum_t \pi_t(j)\ell(q_t^{(j)}, y_t) - \sum_t \pi_t(j)\ell(\sigma(z_j), y_t) + C_2\varepsilon^2 \sum_t \pi_t(j) \\
&\quad \overset{(b)}{\le} C_1 \log T + C_2\varepsilon^2 \sum_t \pi_t(j),
\end{aligned}$$

for some positive constants $C_1, C_2 > 0$, where (a) follows from the definition of the rounding scheme and the upper bound of the rounding error (Lemma 21); (b) holds by that FTL has a regret of $O(\log T)$ under the Brier loss.

Summing this inequality over all $j \in [M]$, we obtain

$$\begin{aligned}
&\sum_{t=1}^T \mathbb{E}_{i\sim A_t\pi_t}\ell(z_i, y_t) - \sum_{t=1}^T \mathbb{E}_{j\sim\pi_t}\ell(\sigma(z_j)) \\
&\quad = \sum_{t=1}^T \sum_{j=1}^M \pi_t(j) \sum_{i=1}^M A_t(i,j)\ell(z_i, y_t) - \sum_{j=1}^M \pi_t(j)\ell(\sigma(z_j), y_t) \\
&\quad = MC_1 \log T + C_2\varepsilon^2 \sum_{t=1}^T \sum_{j=1}^M \pi_t(j) \\
&\quad \le \max\{C_1, C_2\}\left(\frac{\sqrt{K}\log T}{\varepsilon^{K-1}} + \varepsilon^2 T\right),
\end{aligned}$$

where in the last inequality, we use $M = O(\frac{\sqrt{K}}{\varepsilon^{K-1}})$ in Lemma 21. ■

C.3. Algorithm $\mathsf{A}_{\text{lopsided}}$

C.4. Concentration arguments to finish the proof of Theorem 17

We finish the proof of Theorem 17 by upper-bounding the true calibration error using bounds of the pseudo-calibration error. For convenience, denote the pseudo-calibration error under an algorithm A to be $\widetilde{K}_T := \min_{\sigma:\Delta_K\to\Delta_K} \mathbb{E}_{p_t\sim\mathsf{A}}\left[\sum_{t=1}^T \ell(p_t, y_t) - \sum_{t=1}^T \ell(\sigma(p_t), y_t)\right]$. The following lemma is a multiclass extension of Theorem 3 in Luo et al. (2025), relating $\widetilde{K}_T$ to the calibration error.

Lemma 24 *For the Brier loss, for discretization with size M and an algorithm A that always predicts the discretization grid points, with probability at least $1-\delta$ over the randomness in A's predictions $p_1, \ldots, p_T$, we have*

$$K_T(p_{1:T}, y_{1:T}) \le 6\widetilde{K}_T + 96KM \log\frac{4KM}{\delta}.$$

Subroutine 3.2: $\mathsf{A}_{\text{lopsided}}$(Sani et al., 2014)

Input: learning rate $\eta \in (0, \frac{1}{2}]$, initial weights $\{s_1, 1 - s_1\}$
for $t = 1$ **to** T **do**

```
// prediction
```

Output weight $w_t = \frac{s_t}{s_t + 1 - s_1}$.

```
// update
```

Receive loss $\mathbb{E}_{z \sim A_t \pi_t} \ell(z, y_t)$ and $\mathbb{E}_{z \sim b_t} \ell(z, y_t)$ as the losses of experts 1 and 2, respectively.
Compute $\delta_t = g_t^{(2)} - g_t^{(1)}$ and set $s_{t+1} = s_t \cdot (1 + \eta \delta_t)$.
end

Therefore, together with the weaker version proved in Section 5.2 and $M = O(\frac{1}{\varepsilon^{K-1}})$, we have

$$K_T(p_{1:T}, y_{1:T}) \leq O\left(\sqrt{T \log T} + \frac{1}{\varepsilon^{K-1}} \left(K^{1/2} \log T + K^{3/2} \log \frac{4K^{3/2}}{\varepsilon^{K-1}} \right) + \varepsilon^2 T - \log \delta \right)$$
$$= O_{K, \log T}\left(\sqrt{T} + \frac{1}{\varepsilon^{K-1}} \log \frac{1}{\varepsilon} + \varepsilon^2 T \right),$$

with probability at least $1 - \delta$.

C.5. Proof of Lemma 24

We first note the following fact on the closed-form of K_T and $\tilde{K}_T$.

Fact 25 *For the Brier losses, the calibration error under prediction sequence $p_{1:T}$ and outcome sequence $y_{1:T}$ is*

$$K_T = \sum_{p \in \Delta_K} \sum_{t=1}^{T} \mathbf{1}\{p_t = p\} \|p - \rho_T^p\|_2^2 = \sum_{p \in \Delta_K} \sum_{t=1}^{T} \mathbf{1}\{p_t = p\} \sum_{k=1}^{K} \left(p(k) - \rho_T^p(k) \right)^2,$$

where $\rho_T^p = \sum_{t: p_t = p} \frac{y_t}{n_T(p)}$.

The pseudo-calibration error under algorithm A and outcome sequence $y_{1:T}$ is

$$\widetilde{K}_T = \sum_{t=1}^{T} \mathbb{E}_{p \sim P_t} \left[\|p - \tilde{\rho_T}(p)\|^2 \right] = \sum_{t=1}^{T} \sum_{p} P_t(p) \sum_{k=1}^{K} \left(p(k) - \tilde{\rho}_T^p(k) \right)^2,$$

where $P_t(p)$ is the randomized prediction under algorithm A and $\tilde{\rho}_T^p := \frac{\sum_{t=1}^{T} y_t P_t(p)}{\sum_{t=1}^{T} P_t(p)}$.

Denote the ith grid point by z_i in the discretization, and $\rho_T^p(k)$ as $\rho^i(k)$ for simplicity. For any $k \in [K]$, let

$$K_T(k) := \sum_{i \in [M]} \sum_{t=1}^{T} \mathbf{1}\{p_t = z_i\} \left(z_i(k) - \rho^i(k) \right)^2,$$

and

$$\widetilde{K}_T(k) \coloneqq \sum_{i\in[M]} \sum_{t=1}^{T} P_t(z_i)(z_i(k) - \tilde{\rho}^i(k))^2.$$

We have the following lemma.

Lemma 26 *With probability at least* $1-\delta$, $K_T(k) \le 6\widetilde{K}_T(k) + 96M \log \frac{4M}{\delta}$.

Then, with a union bound across $k \in [K]$, with probability at least $1-\delta$, we have

$$K_T = \sum_{k=1}^{K} K_T(k) \le \sum_{k=1}^{K} \left(6\widetilde{K}_T(k) + 96M \log \frac{4MK}{\delta}\right) = 6\widetilde{K}_T + 96MK \log \frac{4MK}{\delta}.$$

C.6. Proof of Lemma 26

Lemma 26 *With probability at least* $1-\delta$, $K_T(k) \le 6\widetilde{K}_T(k) + 96M \log \frac{4M}{\delta}$.

While the proof of Lemma 26 follows almost exactly as the proof of Theorem 3 in Luo et al. (2025), we include it here for completeness.

Proof [Proof of Lemma 26] The proof relies on the following version of Freedman's inequality.

Lemma 27 (Beygelzimer et al. 2011) *Let* $\{X_i\}_{i=1}^n$ *be a martingale difference sequence adapted to the filtration* $\mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_n$*, where* $|X_i| \le B$ *for all* $i \in [n]$*, and* B *is a fixed constant. Define* $\mathcal{V} \coloneqq \sum_{i=1}^n \mathbb{E}\left[X_i^2 \mid \mathcal{F}_{i-1}\right]$*. Then, for any fixed* $\mu \in \left[0, \frac{1}{B}\right], \delta \in [0,1]$*, with probability at least* $1-\delta$*, we have*

$$\left|\sum_{i=1}^{n} X_i\right| \le \mu\mathcal{V} + \frac{\log \frac{2}{\delta}}{\mu}.$$

Fix $i \in [M]$ and define the martingale difference sequence $X_t \coloneqq y_t(k)(P_t(i) - \mathbf{1}\{p_t = z_i\})$ and $Y_t \coloneqq P_t(i) - \mathbf{1}\{p_t = z_i\}$. Observe that $|X_t| \le 1$, $|Y_t| \le 1$ for all t. Fix $\mu_i \in [0,1]$. Applying Lemma 27 to the sequence $X \coloneqq Y \coloneqq X_{1:T}, Y_{1:T}$ and taking a union bound over them, we obtain that with probability at least $1-\delta$,

$$\left|\sum_{t=1}^{T} y_t(k)\left(P_t(i) - \mathbf{1}\left\{p_t = z_i\right\}\right)\right| \le \mu_i\mathcal{V}_X + \frac{\log \frac{4}{\delta}}{\mu_i}, \quad \left|\sum_{t=1}^{T} P_t(i) - \mathbf{1}\left\{p_t = z_i\right\}\right| \le \mu_i\mathcal{V}_Y + \frac{\log \frac{4}{\delta}}{\mu_i}, \tag{8}$$

where $\mathcal{V}_X, \mathcal{V}_Y$ are given by

$$\mathcal{V}_X = \sum_{t=1}^{T} \mathbb{E}\left[X_t^2 \mid \mathcal{F}_{t-1}\right] = \sum_{t=1}^{T} y_t(k) \cdot P_t(i)\left(1 - P_t(i)\right) \le \sum_{t=1}^{T} P_t(i), \text{ and}$$
$$\mathcal{V}_Y = \sum_{t=1}^{T} \mathbb{E}\left[Y_t^2 \mid \mathcal{F}_{t-1}\right] = \sum_{t=1}^{T} P_t(i)\left(1 - P_t(i)\right) \le \sum_{t=1}^{T} P_t(i).$$

The upper tail $\rho^i(k) - \tilde{\rho}^i(k)$ can then be bounded in the following manner:

$$\begin{aligned}
\rho^i(k) - \tilde{\rho}^i(k) &= \frac{\sum_{t=1}^T y_t(k)\mathbf{1}\left\{p_t = z_i\right\}}{\sum_{t=1}^T \mathbf{1}\{p_t = z_i\}} - \frac{\sum_{t=1}^T y_t(k) P_t(i)}{\sum_{t=1}^T P_t(i)} \\
&\overset{(a)}{\leq} \frac{\sum_{t=1}^T y_t(k)\mathbf{1}\{p_t = z_i\}}{\sum_{t=1}^T \mathbf{1}\{p_t = z_i\}} + \frac{\mu_i \sum_{t=1}^T P_t(i) + \frac{\log \frac{4}{\delta}}{\mu_i} - \sum_{t=1}^T y_t(k)\mathbf{1}\{p_t = z_i\}}{\sum_{t=1}^T P_t(i)} \\
&= \frac{\sum_{t=1}^T y_t(k)\mathbf{1}\{p_t = z_i\}}{\left(\sum_{t=1}^T \mathbf{1}\{p_t = z_i\}\right)\left(\sum_{t=1}^T P_t(i)\right)} \cdot \left(\sum_{t=1}^T P_t(i) - \sum_{t=1}^T \mathbf{1}\{p_t = z_i\}\right) \\
&\quad + \frac{\mu_i \sum_{t=1}^T P_t(i) + \frac{\log \frac{4}{\delta}}{\mu_i}}{\sum_{t=1}^T P_t(i)} \\
&\overset{(b)}{\leq} \frac{\sum_{t=1}^T y_t(k)\mathbf{1}\{p_t = z_i\}}{\left(\sum_{t=1}^T \mathbf{1}\{p_t = z_i\}\right)\left(\sum_{t=1}^T P_t(i)\right)} \cdot \left(\mu_i \sum_{t=1}^T P_t(i) + \frac{\log \frac{4}{\delta}}{\mu_i}\right) \\
&\quad + \frac{\mu_i \sum_{t=1}^T P_t(i) + \frac{\log \frac{4}{\delta}}{\mu_i}}{\sum_{t=1}^T P_t(i)} \\
&\overset{(c)}{\leq} 2\mu_i + \frac{2\log \frac{4}{\delta}}{\mu_i \sum_{t=1}^T P_t(i)},
\end{aligned}$$

where (a) and (b) follow from (8), and (c) follows by $y_t(k)\mathbf{1}\{p_t = z_i\} \leq \mathbf{1}\{p_t = z_i\}$. The lower tail can be bounded in an exact manner as

$$\begin{aligned}
\tilde{\rho}^i(k) - \rho^i(k) &= \frac{\sum_{t=1}^T y_t(k) P_t(i)}{\sum_{t=1}^T P_t(i)} - \frac{\sum_{t=1}^T y_t(k)\mathbf{1}\{p_t = z_i\}}{\sum_{t=1}^T \mathbf{1}\{p_t = z_i\}} \\
&\leq \frac{\sum_{t=1}^T y_t(k)\mathbf{1}\{p_t = z_i\} + \mu_i \sum_{t=1}^T P_t(i) + \frac{\log \frac{4}{\delta}}{\mu_i}}{\sum_{t=1}^T P_t(i)} - \frac{\sum_{t=1}^T y_t(k)\mathbf{1}\{p_t = z_i\}}{\sum_{t=1}^T \mathbf{1}\{p_t = z_i\}} \\
&= \frac{\sum_{t=1}^T y_t(k)\mathbf{1}\{p_t = z_i\}}{\left(\sum_{t=1}^T P_t(i)\right)\left(\sum_{t=1}^T \mathbf{1}\{p_t = z_i\}\right)} \cdot \left(\sum_{t=1}^T \mathbf{1}\{p_t = z_i\} - \sum_{t=1}^T P_t(i)\right) \\
&\quad + \frac{\mu_i \sum_{t=1}^T P_t(i) + \frac{\log \frac{4}{\delta}}{\mu_i}}{\sum_{t=1}^T P_t(i)} \\
&\leq \frac{\sum_{t=1}^T y_t(k)\mathbf{1}\{p_t = z_i\}}{\left(\sum_{t=1}^T \mathbf{1}\{p_t = z_i\}\right)\left(\sum_{t=1}^T P_t(i)\right)} \cdot \left(\mu_i \sum_{t=1}^T P_t(i) + \frac{\log \frac{4}{\delta}}{\mu_i}\right) \\
&\quad + \frac{\mu_i \sum_{t=1}^T P_t(i) + \frac{\log \frac{4}{\delta}}{\mu_i}}{\sum_{t=1}^T P_t(i)} \\
&\leq 2\mu_i + \frac{2\log \frac{4}{\delta}}{\mu_i \sum_{t=1}^T P_t(i)}.
\end{aligned}$$

Combining both bounds, we have shown that for a fixed $\mu_i \in [0,1]$, $|\rho^i(k) - \tilde{\rho}^i(k)| \le 2\mu_i + \frac{\log\frac{4}{\delta}}{\mu_i \sum_{t=1}^T P_t(i)}$ holds with probability at least $1-\delta$. Taking a union bound over all i, with probability $1-\delta$, the following holds simultaneously for all i,

$$\left|\sum_{t=1}^T y_t(k)\left(P_t(i) - \mathbf{1}\{p_t = z_i\}\right)\right| \le \mu_i \sum_{t=1}^T P_t(i) + \frac{\log\frac{4M}{\delta}}{\mu_i}, \tag{9}$$

$$\left|\sum_{t=1}^T P_t(i) - \mathbf{1}\{p_t = z_i\}\right| \le \mu_i \sum_{t=1}^T P_t(i) + \frac{\log\frac{4M}{\delta}}{\mu_i}, \tag{10}$$

$$\left|\rho^i(k) - \tilde{\rho}^i(k)\right| \le 2\mu_i + \frac{2\log\frac{4M}{\delta}}{\mu_i \sum_{t=1}^T P_t(i)}. \tag{11}$$

Consider the function $g(\mu) := \mu + \frac{a}{\mu}$, where $a \ge 0$ is a fixed constant. Clearly, $\min_{\mu\in[0,1]} g(\mu) = 2\sqrt{a}$ when $a \le 1$, and $1 + a$ otherwise. Minimizing the bound in (11) with respect to μ_i, we obtain

$$\left|\rho^i(k) - \tilde{\rho}^i(k)\right| \le \begin{cases} 4\sqrt{\frac{\log\frac{4M}{\delta}}{\sum_{t=1}^T P_t(i)}}, \text{ when } \log\frac{4M}{\delta} \le \sum_{t=1}^T P_t(i), \\ 2 + \frac{2\log\frac{4M}{\delta}}{\sum_{t=1}^T P_t(i)}, \text{ when } \log\frac{4M}{\delta} > \sum_{t=1}^T P_t(i). \end{cases}$$

Therefore, when $\sum_{t=1}^T P_t(i)$ is tiny, which is possible if algorithm A does not allocate enough probability mass to the index i, the bound obtained is large making it much worse than the trivial bound $\left|\rho^i(k) - \tilde{\rho}^i(k)\right| \le 1$ which follows since $\rho^i(k), \tilde{\rho}^i(k) \in [0,1]$ by definition. Based on this reasoning, we define the set

$$\mathcal{I} := \left\{ i \in [M], \text{ s.t. } \log\frac{4M}{\delta} \le \sum_{t=1}^T P_t(i) \right\}, \tag{12}$$

and let $\bar{\mathcal{I}} := [M] \setminus \mathcal{I}$. We bound $(\rho^i(k) - \tilde{\rho}^i(k))^2$ as

$$\left(\rho^i(k) - \tilde{\rho}^i(k)\right)^2 \le \begin{cases} \frac{16\log\frac{4M}{\delta}}{\sum_{t=1}^T P_t(i)} & \text{if } i \in \mathcal{I}, \\ 1 & \text{otherwise.} \end{cases} \tag{13}$$

Similarly, $\left|\sum_{t=1}^T P_t(i) - \mathbf{1}\{p_t = z_i\}\right|$ can be bounded by substituting the optimal μ_i obtained above in (10); we obtain

$$\left|\sum_{t=1}^T P_t(i) - \mathbf{1}\{p_t = z_i\}\right| \le \begin{cases} 2\sqrt{\log\frac{4M}{\delta} \sum_{t=1}^T P_t(i)} & \text{if } i \in \mathcal{I} \\ \sum_{t=1}^T P_t(i) + \log\frac{4M}{\delta} & \text{otherwise.} \end{cases} \tag{14}$$

Equipped with (13) and (14), we proceed to bound $K_T(k)$ in the following manner:

$$\begin{aligned} K_T(k) &= \sum_{i\in[M]} \sum_{t=1}^T \mathbf{1}\{p_t = z_i\} \left(z_i(k) - \rho^i(k)\right)^2 \\ &\le 2 \sum_{i\in[M]} \sum_{t=1}^T \mathbf{1}\{p_t = z_i\} \left(\left(z_i(k) - \tilde{\rho}^i(k)\right)^2 + \left(\rho^i(k) - \tilde{\rho}^i(k)\right)^2\right), \end{aligned}$$

where the inequality holds because $(a+b)^2 \le 2a^2 + 2b^2$ for all $a, b \in \mathbb{R}$. To further bound the term above, we split the summation into two terms $\mathcal{T}_1, \mathcal{T}_2$ defined as

$$\mathcal{T}_1 := \sum_{i \in \mathcal{I}} \sum_{t=1}^{T} \mathbf{1}\{p_t = z_i\} \left(\left(z_i(k) - \tilde{\rho}^i(k)\right)^2 + \left(\rho^i(k) - \tilde{\rho}^i(k)\right)^2 \right),$$

$$\mathcal{T}_2 = \sum_{i \in \bar{\mathcal{I}}} \sum_{t=1}^{T} \mathbf{1}\{p_t = z_i\} \left(\left(z_i(k) - \tilde{\rho}^i(k)\right)^2 + \left(\rho^i(k) - \tilde{\rho}^i(k)\right)^2 \right),$$

and bound $\mathcal{T}_1$ and $\mathcal{T}_2$ individually. We bound $\mathcal{T}_1$ as

$$\begin{aligned}
\mathcal{T}_1 &\overset{(a)}{\le} \sum_{i \in \mathcal{I}} \left(\sum_{t=1}^{T} P_t(i) + 2\sqrt{\log \frac{4M}{\delta} \sum_{\tau=1}^{T} P_\tau(i)} \right) \left(\left(z_i(k) - \tilde{\rho}^i(k)\right)^2 + \frac{16 \log \frac{4M}{\delta}}{\sum_{\tau=1}^{T} P_\tau(i)} \right) \\
&= \sum_{i \in \mathcal{I}} \sum_{t=1}^{T} P_t(i) \left(z_i(k) - \tilde{\rho}^i(k)\right)^2 + 16 \log \frac{4M}{\delta} |\mathcal{I}| + \\
&\quad 2 \sum_{i \in \mathcal{I}} \sqrt{\log \frac{4M}{\delta} \sum_{\tau=1}^{T} P_\tau(i)} \left(\left(z_i(k) - \tilde{\rho}^i(k)\right)^2 + \frac{16 \log \frac{4M}{\delta}}{\sum_{\tau=1}^{T} P_\tau(i)} \right) \\
&\overset{(b)}{\le} \sum_{i \in \mathcal{I}} \sum_{t=1}^{T} P_t(i) \left(z_i(k) - \tilde{\rho}^i(k)\right)^2 + 16 \log \frac{4M}{\delta} |\mathcal{I}| + \\
&\quad 2 \sum_{i \in \mathcal{I}} \sum_{\tau=1}^{T} P_\tau(i) \left(\left(z_i(k) - \tilde{\rho}^i(k)\right)^2 + \frac{16 \log \frac{4M}{\delta}}{\sum_{\tau=1}^{T} P_\tau(i)} \right) \\
&= 3 \sum_{i \in \mathcal{I}} \sum_{t=1}^{T} P_t(i) \left(z_i(k) - \tilde{\rho}^i(k)\right)^2 + 48 \log \frac{4M}{\delta} |\mathcal{I}|
\end{aligned}$$

where (a) follows by substituting the bounds from (13) and (14); while (b) follows since by the definition of $\mathcal{I}$ in (12), we have $\sqrt{\log \frac{4M}{\delta} \sum_{\tau=1}^{T} P_\tau(i)} \le \sum_{\tau=1}^{T} P_\tau(i)$. Next, we bound $\mathcal{T}_2$ as

$$\begin{aligned}
\mathcal{T}_2 &\overset{(a)}{\le} \sum_{i \in \overline{\mathcal{I}}} \left(2 \sum_{t=1}^{T} P_t(i) + \log \frac{4M}{\delta} \right) \left(\left(z_i(k) - \tilde{\rho}^i(k)\right)^2 + 1 \right) \\
&\overset{(b)}{\le} 2 \sum_{i \in \overline{\mathcal{I}}} \sum_{t=1}^{T} P_t(i) \left(z_i(k) - \tilde{\rho}^i(k)\right)^2 + 2 \sum_{i \in \overline{\mathcal{I}}} \sum_{t=1}^{T} P_t(i) + 2 \log \frac{4M}{\delta} |\overline{\mathcal{I}}| \\
&\overset{(c)}{\le} 2 \sum_{i \in \overline{\mathcal{I}}} \sum_{t=1}^{T} P_t(i) \left(z_i(k) - \tilde{\rho}^i(k)\right)^2 + 4 \log \frac{4M}{\delta} |\overline{\mathcal{I}}|
\end{aligned}$$

where (a) follows by substituting the bounds from (13) and (14); (b) follows by bounding $(z_i(k) - \tilde{\rho}^i(k))^2 \le 1$; and (c) follows from the definition of $\mathcal{I}$ in (12). Collecting the bounds on $\mathcal{T}_1$ and $\mathcal{T}_2$, we obtain

$$\begin{aligned}\mathcal{T}_1 + \mathcal{T}_2 &\leq 3 \sum_{i\in[M]} \sum_{t=1}^{T} P_t(i) \left(z_i(k) - \tilde{\rho}^i(k)\right)^2 + 48 \log \frac{4M}{\delta} |\mathcal{I}| + 4 \log \frac{4M}{\delta} |\overline{\mathcal{I}}| \\ &\leq 3\widetilde{K}_T(k) + 48M \log \frac{4M}{\delta},\end{aligned}$$

where the last inequality follows from the definition of $\widetilde{K}_T(k)$ and since $|\mathcal{I}| + |\overline{\mathcal{I}}| = M$. Since $K_T(k) \leq 2\left(\mathcal{T}_1 + \mathcal{T}_2\right)$, we have shown that

$$K_T(k) \leq 6\widetilde{K}_T(k) + 96M \log \frac{4M}{\delta},$$

with probability at least $1 - \delta$. This completes the proof. ■