arXiv:2302.10863v1 [cs.LG] 21 Feb 2023

# A Unifying Perspective on Multi-Calibration: Unleashing Game Dynamics for Multi-Objective Learning*

Nika Haghtalab, Michael I. Jordan, and Eric Zhao

University of California, Berkeley
{nika,jordan,eric.zh}@berkeley.edu

**Abstract**

We provide a unifying framework for the design and analysis of multi-calibrated and moment-multi-calibrated predictors. Placing the multi-calibration problem in the general setting of *multi-objective learning*—where learning guarantees must hold simultaneously over a set of distributions and loss functions—we exploit connections to game dynamics to obtain state-of-the-art guarantees for a diverse set of multi-calibration learning problems. In addition to shedding light on existing multi-calibration guarantees, and greatly simplifying their analysis, our approach yields a $1/\varepsilon^2$ improvement in the number of oracle calls compared to the state-of-the-art algorithm of Jung et al. [19] for learning deterministic moment-calibrated predictors and an exponential improvement in $k$ compared to the state-of-the-art algorithm of Gopalan et al. [14] for learning a $k$-class multi-calibrated predictor. Beyond multi-calibration, we use these game dynamics to address existing and emerging considerations in the study of group fairness and multi-distribution learning.

## 1 Introduction

In recent years, multi-calibration has emerged as a powerful tool for addressing fairness considerations in machine learning. Founded on the principle of calibrated forecasting [5, 10]—which requires that among instances that are assigned prediction probability $h(x) = v$, a fraction $v$ of those instances demonstrate a positive outcome—multi-calibration seeks to provide calibration guarantees at finer and richer levels of decision making. These include the multi-calibration notion of Hebert-Johnson et al. [18] that additionally seeks calibration across large and possibly overlapping collection of sub-populations, the multi-class multi-calibration notion of Gopalan et al. [14] that considers calibration for predictors with rich label sets, and the moment multi-calibration notion of Jung et al. [19] that in addition to average outcome seeks calibrated estimates of higher moments. Addressing these additional considerations in calibration has contributed to the versatility of multi-calibration and its connections to fairness, robust accuracy guarantees, computational indistinguishability, and conformal predictions [18, 19, 13, 15, 6].

This development of new applications and variants of multi-calibration has introduced a plethora of algorithms—with an accompanying lack of clarity as to which algorithms are appropriate for which problem settings, particularly novel problem settings that have emerged recently. There have been promising attempts to unify these approaches, such as outcome indistinguishability [6], but

*Authors are ordered alphabetically.

| Problem | Complexity | Previous Results | Our Results | Reference |
|---|---|---|---|---|
| Moment MC | Oracle | $O\left(r\varepsilon^{-4}\right)$ [19] | $O\left(r\varepsilon^{-2}\right)$ | Corr 4.12 |
| | Sample | $\widetilde{O}\left(\sqrt{r}\varepsilon^{-4}(d+\ln(1/\lambda))\right)$ [19] | $\widetilde{O}\left(\sqrt{r}\varepsilon^{-3}(d+\ln(1/\lambda))\right)$ | |
| $k$-class MC | Oracle | $O\left(k\varepsilon^{-2}\right)$ [14] | $O\left(\ln(k)\varepsilon^{-2}\right)$ | Corr 4.7 |
| | Sample | $\widetilde{O}\left(\sqrt{k}\varepsilon^{-3}(d+k\ln(1/\lambda))\right)$ [14] | $\widetilde{O}\left(\sqrt{\ln(k)}\varepsilon^{-3}(d+k\ln(1/\lambda))\right)$ | |
| Moment ND MC | Sample | $\widetilde{O}\left((d+\ln(r/\lambda))\varepsilon^{-2}\right)$ [15] | $O\left((d+\ln(r/\lambda))\varepsilon^{-2}\right)$ | Corr 4.13 |
| $k$-class ND MC | Sample | $\widetilde{O}\left((d+k\ln(1/\lambda))\varepsilon^{-2}\right)$ [15] | $\widetilde{O}\left((d+k\ln(1/\lambda))\varepsilon^{-2}\right)$ | Corr 4.5 |

Table 1: This table summarizes the oracle and sample complexity rates we obtain for various multi-calibration (MC) problems, compared against the previously best known rates. We let $\lambda$ denote calibration bin size, $\mathcal{S}$ the set of groups one wants to be calibrated on, $d$ the VC dimension of set $\mathcal{S}$, $k$ the number of classes, and $r$ the number of higher-order moments to be calibrated. The bottom two rows concern relaxed settings where non-deterministic (ND) solutions are acceptable and for which we provide matching rates.

they have had limited success in providing a simple overarching conceptual framework that can capture the diverse needs of calibration. In this paper, we tackle this challenge head-on, developing a basic and versatile algorithmic framework that guides the design of simple and efficient multi-calibrated learning algorithms for a wide range of increasingly specialized and diverse considerations in calibration.

Our approach builds on well-established game dynamics for learning problems [see, e.g., 11, 8]. We show that most results in multi-calibration can be formulated in terms of one of two classical scenarios in the dynamical formulation of two-player zero-sum games: no-regret versus no-regret and no-regret versus best-response. Plugging in different no-regret and best-response algorithms results in different multi-calibration algorithms exhibiting different trade-offs, obviating the need for highly individualized, detailed, and careful analysis of specific settings. We also present a no-regret algorithm that leverages the structure of any calibration-like function class to obtain improved sample-complexity and oracle-complexity for several multi-calibration problems, including mean-conditioned moment multi-calibrations and multi-class multi-calibration.

Still more generally, we introduce *multi-objective learning* as a framework for obtaining learning guarantees that hold simultaneously over a collection of distributions and loss functions. Beyond multi-calibration, this framework readily accommodates existing frameworks for collaborative learning [3, 24], multi-distribution learning [16], agnostic federated learning [23], and a variety of group fairness frameworks [27, 26, 28]. Once again, we use the versatility of the game dynamics in multi-objective learning problems to improve upon several existing results in these literatures.

### 1.1 Overview of Results

**Simple and unifying framework.** In Section 3, we give a general overview of game dynamics as a concise and unifying framework for obtaining multi-objective learning guarantees. We apply these techniques to several problems in multi-calibration by introducing a general-purpose no-regret algorithm (Lemma 4.2) for calibration-like objectives. In addition to simplifying past techniques and providing improved learning guarantees, our framework takes a step towards unifying seemingly different approaches used in the field.

**Improved guarantees.** We show that our classical learning-in-games dynamics framework provides the best-known sample complexity rates for most multi-calibration objectives. In particular, for the problem of $k$-class multi-calibration (for deterministic predictors) we obtain an exponential

improvement in $k$ over Gopalan et al. [14], by using only $O\left(\log(k)/\varepsilon^2\right)$ rather than $O\left(k/\varepsilon^2\right)$ calls to an agnostic oracle. For moment multi-calibration we use $O\left(1/\varepsilon^3\right)$ samples or $O\left(1/\varepsilon^2\right)$ calls to an oracle to find a deterministic moment multi-calibrated predictor. This improves the state-of-the-art results of Jung et al. [19] who uses $O\left(1/\varepsilon^4\right)$ samples or $O\left(1/\varepsilon^4\right)$ calls to an oracle. To achieve these improvements, we design a no-regret algorithm in Lemma 4.2 that works for any calibration-like functions—one that is fundamental to all applications of multi-calibration.

**Simplified results and new considerations.** In addition to the above improvements, we broadly match the best-known guarantees of most multi-calibrated settings while obviating the need for their individualized treatment, including for non-deterministic moment calibration and multi-calibration results of Gupta et al. [15], Noarov et al. [25] in Corollaries 4.13 and 4.5 respectively. We also show our game dynamics analysis extends to new variants of multi-calibration, such as in agnostic settings, and beyond multi-calibration, such as for multi-group fairness (matching [28]).

### 1.2 Technical Overview

The design of our algorithms is based on the construction of game dynamics—pitting players against one another across many iterations—on game representations of multi-objective learning. It is well-established that when the players in these dynamics use no-regret strategies, their time-averaged actions give a randomized predictor that converges to an optimal solution. When one of the players is best responding, we show that it may even be possible to extract a deterministic solution from these dynamics (Lemma 3.2). One implementation of multi-objective learning game dynamics is to have each player run an online learning algorithm, like Hedge, on the empirical game—since their losses form a martingale, their true regret will quickly converge to their empirical regret (e.g. see Haghtalab et al. [Lemma B.1 16]). However, convergence may be slow for players tasked with optimizing over candidate hypotheses, as standard online learning algorithms would give them regret bounds scaling in the complexity of $\mathcal{H}$ (or $\mathcal{X}$).

In this regard, multi-calibration is distinguished from the general setting of multi-objective learning by just two properties that are both rooted in the special linear structure of the calibration loss. The first is a folklore property that has been recently formalized by Hart [17] establishing that the learner can compute her best response in a distribution-free way (without observing samples). This means that there are game dynamics for calibration problems that converge faster than is typically possible: at a rate independent of the complexity of one's domain $|\mathcal{X}|$. The second property that we formalize in Lemma 4.2 is that the learner has a no-regret online learning strategy that does not use randomization. The latter property has been implicitly used in multi-calibration literature for specific choices of online algorithms. On the other hand, our Lemma 4.2 makes this explicit and shows that any no-regret algorithm over a finite set can be used as a basis for creating a deterministic no-regret algorithm for the calibration loss. By making this explicit, we identify opportunities for unifying and simplifying existing algorithms and improving multi-calibration sample complexity rates.

### 1.3 Related Works

**Multi-calibration and related notions.** The study of calibration in statistics took its roots in forecasting [5, 17]. The classical literature also included work on calibration across multiple sub-populations [9]. Motivated by fairness considerations, a formal definition of multi-calibration was presented by Hebert-Johnson et al. [18], and that formulation has found a wide range of applications

and conceptual connections to Bayes optimality, conformal predictions, and computational indistinguishability [18, 19, 13, 15, 6, 20]. Algorithms for multi-calibration have largely developed along two lines, one studying oracle-efficient boosting-like algorithms [see, e.g., 18, 22, 14, 6] and another studying algorithms with flavors of online optimization [see, e.g., 15, 25]. Our work establishes that the contrast between these lines of work are entirely attributable to different choices of game dynamics.

**Other multi-objective learning problems.** More broadly, multi-objective and multi-distribution learning have been studied in settings involving fairness, collaboration, and robustness. Blum et al. [3] initiated the study of optimal sample complexity for learning predictors with near-optimal accuracy across multiple populations. The recent work of Haghtalab et al. [16] uses game dynamics to obtain tight sample complexity bounds (including agnostic and well-specified settings), improving over past work that similarly drew on boosting-like and online optimization techniques [4, 24, 23]. Related notions of group fairness have also been studied by Kearns et al. [21] with a focus on using weak oracles for optimization. More recently, multi-objective learning has been viewed as a middle ground between settings where sub-populations are mutually compatible with the best known bounds arising by reduction to the online learning framework of sleeping experts [26, 28, 2]. In Section 5, we show how our framework gives matching (or improved) guarantees relative to these literature.

## 2 Preliminaries

We will use $\mathcal{X}$ and $\mathcal{Y}$ to denote the feature space and label space of our learning problems. Two common label sets we work with are $\mathcal{Y} \coloneqq \{1, \dots, k\}$ and $\mathcal{Y} \coloneqq [0,1]^k$, using the latter especially for calibration with $k$ classes. We consider data distributions $D$ over the set $\mathcal{Z} \coloneqq \mathcal{X} \times \mathcal{Y}$. We consider a hypothesis (or predictor) class $\mathcal{H}$ that is a set of functions mapping $\mathcal{X}$ to labels $\mathcal{Y}$. We consider loss functions of the form $\ell : \mathcal{H} \times \mathcal{Z} \to [0,1]$. Given a data distribution $D$ and a loss function $\ell$, we denote the expected loss of any hypothesis $h$ by $\mathcal{L}_{D,\ell}(h) \coloneqq \mathbb{E}_{(x,y)\sim D}\left[\ell\left(h,(x,y)\right)\right]$ and suppress subscripts when clear from context. We also consider non-deterministic choices $p \in \Delta\mathcal{H}$ and $q \in \Delta\mathcal{D}$, where $\Delta$ denotes a simplex of probability distributions over a base class. In this case, we define their expected loss by $\mathcal{L}_{q,\ell}(p) \coloneqq \mathbb{E}_{D\sim q, h\sim p}\left[\mathcal{L}_{D,\ell}(h)\right]$.

**Multi-objective learning.** Multi-objective learning is concerned with finding a hypothesis (or a distribution over hypotheses) that performs well over any loss function and any distribution, from a predetermined set of losses and distributions. A *multi-objective learning* instance is defined by $(D, \mathcal{G}, \mathcal{H})$, where $\mathcal{H}$ is a hypothesis class, $\mathcal{G}$ is a set of loss functions $\ell : (\mathcal{H} \times \mathcal{Z}) \to [0,1]$, and $D$ is a data distribution. The goal of multi-objective learning is to return a hypothesis $h$, or more generally a distribution $p \in \Delta\mathcal{H}$, such that

$$\max_{\ell\in\mathcal{G}} \mathcal{L}_{D,\ell}(p) \leq \min_{h^*\in\mathcal{H}} \max_{\ell\in\mathcal{G}} \mathcal{L}_{D,\ell}(h^*) + \varepsilon. \tag{1}$$

We call $\mathcal{L}^*(h) \coloneqq \max_{\ell\in\mathcal{G}} \mathcal{L}_{D,\ell}(h)$ and the analogously defined $\mathcal{L}^*(p)$ the *multi-objective loss*. We call a solution $p$ that meets Equation (1) an *$\varepsilon$-optimal solution* to the multi-objective instance $(D, \mathcal{G}, \mathcal{H})$. Note that $p$ can take many forms—a non-deterministic hypothesis or a deterministic hypothesis. Furthermore, a deterministic solution $h$ may be *proper* (i.e., $h \in \mathcal{H}$) or *improper*. While we prefer to find a deterministic solution, depending on the application, non-deterministic solutions may also be useful.

**Oracles.** We access any distribution $D \in \mathcal{D}$ using only an *example oracle*, $\mathrm{EX}_D$, that samples $z \sim D$. We generally work with known loss functions $\mathcal{G}$, though the class may be large. In some cases, we interact with $\mathcal{G}$ only through optimization oracles. A $(c, \varepsilon)$-agnostic learning oracle for class $\mathcal{G}$ (and distribution $D$) returns a $q \in \Delta\mathcal{G}$ given any $h \in \mathcal{H}$, such that $\mathcal{L}_{D,q}(h) + \varepsilon \geq c \cdot \max_{\ell^* \in \mathcal{G}} \mathcal{L}_{D,\ell^*}(h)$. We note that such learning oracles are readily available for finite $\mathcal{G}$ (such as the ERM oracle) and structured infinite classes, using a small number of example oracle calls to $D$. We also consider oracles return solutions that directly compete with the minmax value, instead of being instance-wise nearly optimal. That is, we define a *weak* $\varepsilon$-oracle that takes as input a hypothesis $h \in \mathcal{H}$ and returns a loss $\ell \in \mathcal{G}$ such that $\mathcal{L}_{D,\ell}(h) \geq \varepsilon + \min_{h^* \in \mathcal{H}} \max_{\ell^* \in \mathcal{G}} \mathcal{L}_{D,\ell^*}(h^*)$ if there exists a choice of $\ell'$ where $\mathcal{L}_{D,\ell'}(h) \geq 2\varepsilon + \min_{h^*} \max_{\ell^*} \mathcal{L}_{D,\ell^*}(h^*)$, i.e., if $\ell'$ beat the minmax value by $2\varepsilon$. We call a learning algorithm *oracle-efficient* if it interacts with $\mathcal{G}$ only through one of these oracles. We generally state our results in terms of $(1, \varepsilon)$-agnostic learning oracles but all of our results extend to weak $\varepsilon$-oracles as well (see Remark 3.3). Since we can always find a covering of $\mathcal{G}$ or just interact with $\mathcal{G}$ through an oracle, we will use $|\ln(\mathcal{G})|$ instead of VC dimension $\mathrm{VC}(\mathcal{G})$, though our results hold for both with at most an extra $\ln(1/\varepsilon)$ factor.

**Games Formulation and Dynamics.** Multi-objective learning is closely related to two-player zero-sum games. For ease of notation, we now briefly consider an abstract two-player zero-sum game, where $A_i$ is the action set of player $i$, $\phi_i : \prod_j A_j \to [0, 1]$ is the loss of player $i$, and $\phi_1(a) = -\phi_2(a)$ for all strategy profiles $a \in \prod_j A_j$. An $\varepsilon$-dominant strategy for player $i$ is a pure strategy $a_i$ such that for all $a_{-i} \in A_{-i}$ and $a_i' \in A_i$, $\phi_i(a_i, a_{-i}) \leq \phi(a_i', a_{-i}) + \varepsilon$. For any sequence of played actions, $a^{(1)}, \ldots, a^{(T)}$, the regret of player $i$ is defined as,

$$\mathrm{Reg}_i(a^{(1:T)}) \coloneqq \sum_{t=1}^{T} \phi_i(a^{(t)}) - \min_{a_i' \in A_i} \sum_{t=1}^{T} \phi_i(a_{-i}^{(t)}, a_i').$$

We also introduce a *weak-regret* notion for player $i$:

$$\mathrm{WReg}_i(a^{(1:T)}) \coloneqq \sum_{t=1}^{T} \phi_i(a^{(t)}) - T \cdot \min_{a_i' \in A_i} \max_{a_{-i}' \in A_{-i}} \phi_i(a').$$

We define regret and weak regret analogously for sequences of mixed strategies. Formally, we consider a multi-objective setting as a (stochastic) two-player zero-sum game where the learner's action set is $A_1 = \mathcal{H}$, the adversary's action set is $A_2 = \mathcal{G}$, and $\phi_1(h, \ell) = \mathcal{L}_{D,\ell}(h)$. We denote the regret (or weak regret) of learner $\mathbf{L}$ and adversary $\mathbf{A}$ with $\mathrm{Reg}_{\mathbf{L}}$ and $\mathrm{Reg}_{\mathbf{A}}$, respectively (and $\mathrm{WReg}_{\mathbf{L}}$ and $\mathrm{WReg}_{\mathbf{A}}$, respectively).

Since distributions in $\mathcal{D}$ can only be accessed via example oracles, the players do not directly observe the value of $\phi(h, (D, \ell))$; rather, they estimate it via random samples $z \sim D$. Therefore, the history of play for multi-objective problems includes both the chosen strategies of players $\left\{p^{(t)}, q^{(t)}\right\}_{t=1}^{T}$ as well as random samples $z^{(t)} \sim D$. We define the corresponding notions of *empirical* regret and weak regret, where player's losses are their empirical loss on the realized samples $z^{(t)}$, as follows

$$\widehat{\mathrm{Reg}}_{\mathbf{A}}\left(\left\{p^{(t)}, q^{(t)}, z^{(t)}\right\}_{t=1}^{T}\right) \coloneqq \max_{\ell^* \in \mathcal{G}} \sum_{t=1}^{T} \ell^*(p^{(t)}, z^{(t)}) - \sum_{t=1}^{T} q^{(t)}(h^{(t)}, z^{(t)})$$

$$\widehat{\mathrm{WReg}}_{\mathbf{L}}\left(\left\{p^{(t)}, q^{(t)}, z^{(t)}\right\}_{t=1}^{T}\right) \coloneqq \sum_{t=1}^{T} q^{(t)}(p^{(t)}, z^{(t)}) - \min_{h^* \in \mathcal{H}} \max_{\ell^* \in \mathcal{G}} \sum_{t=1}^{T} \ell^*(h^*, z^{(t)}).$$

We note that standard no-regret algorithms, such as Hedge [12], can be used to bound the empirical regret of the adversary by $O(\sqrt{T\log(|\mathcal{G}|)})$ and the empirical regret of the learner by $O(\sqrt{T\log(|\mathcal{H}|)})$. Later in this paper, we give tighter empirical regret bounds for special function classes.

An important feature of empirical regret is that, as long as learner's and adversary's choices $h^{(t)}, \ell^{(t)}$ are made before $z^{(t)}$ is realized, the empirical and true regret (respectively weak regret) are close to each other. The next lemma, which we will state from the perspective of the adversary, formalizes this more generally for martingale sequences.

**Lemma 2.1.** *Let $\ell^{(1:T)}$ be the result of running an online learning algorithm on an adversarially chosen sequence $h^{(1:T)}$ which is constrained to be a martingale with respect to $z^{(1:T)}$. Then, with probability at least $1-\delta$,*

$$Reg_{\mathbf{A}}(\ell^{(1:T)}, h^{(1:T)}) \leq \mathbb{E}\left[\widehat{Reg}_{\mathbf{A}}\left(\ell^{(1:T)}, h^{(1:T)}, z^{(1:T)}\right)\right] + O\left(\sqrt{T\log(|\mathcal{H}|/\delta)}\right).$$

# 3 Multi-Objective Learning with Game Dynamics

Expressing multi-objective learning as a game allows one to draw on learning-in-games techniques for designing algorithms. We overview two such techniques, each relating to a game dynamic.

**No-regret vs. no-regret (NRNR) dynamics.** Game dynamics where two players each use a no-regret algorithm to choose their actions have long played a role in empirical convergence to notions of equilibria [12]. Generally, these dynamics will lead to non-deterministic near-optimal solutions. In the following lemma, we slightly generalize these dynamics to a case where the adversary is no-regret and the learner only has *no weak-regret*, i.e., the learner only needs to be on-average happier than they would be at the true equilibrium.

**Lemma 3.1** (No-Regret vs. No-Regret)**.** *Let $(p,q)^{(1:T)}$ be a play history between a learner $\mathbf{L}$ with actions $\mathcal{H}$ and an adversary $\mathbf{A}$ with actions $\mathcal{G}$. If $T\varepsilon \geq WReg_{\mathbf{L}}\left((p,q)^{(1:T)}\right)$ and $T\cdot\varepsilon \geq Reg_{\mathbf{A}}\left((p,q)^{(1:T)}\right)$, then a uniform distribution on $p^{(1:T)}$ is a $2\varepsilon$-optimal solution for the multi-objective learning problem $(\mathcal{G},\mathcal{H})$.*

*Proof.* Since the adversary is no-regret (left inequality) and the learner is no-weak-regret (right),

$$\max_{\ell^*\in\mathcal{G}} \sum_{t=1}^{T} \mathcal{L}^*(p^{(t)}) - T\varepsilon \leq \sum_{t=1}^{T} \mathcal{L}^{(t)}(p^{(t)}), \quad \sum_{t=1}^{T} \mathcal{L}^{(t)}(p^{(t)}) \leq T\varepsilon + T\min_{h^*\in\mathcal{H}}\max_{\ell^*\in\mathcal{G}} \mathcal{L}^*(h^*),$$

the claim immediately follows

$$\max_{\ell^*\in\mathcal{G}} \frac{1}{T}\sum_{t=1}^{T} \mathcal{L}^*(p^{(t)}) \leq 2\varepsilon + \min_{h^*\in\mathcal{H}}\max_{\ell^*\in\mathcal{G}} \mathcal{L}^*(h^*) = 2\varepsilon + \mathrm{OPT}.$$

□

Generalizing these dynamics to weak regret allows us to leverage algorithms that are known to be capable of beating the minmax optimal, rather than always being no-regret. We give one such algorithm for calibration-like objectives in the next section. No-regret dynamics are also powerful because they are robust to noise, as seen in Lemma 2.1; for instance, when we implement our adversaries with Hedge, Lemma 2.1 allows us to directly work with distributional rather than empirical regret.

**No-regret vs. best-response (NRBR) dynamics.** NRBR are a form of no-regret dynamics where one player must also (on-average) best-respond. This comes at a cost—best-responding often requires greater sample complexity than being no-regret. While NRBR ensures that the empirical play converges to equilibria, it also ensures that the no-regret player plays a reasonably good strategy (beating the pure strategy minmax) at some time step. This means, for instance, that we can recover a *deterministic* approximate solution for the no-regret player if they only play pure strategies at each timestep.

**Lemma 3.2** (No-Regret vs. Best-Response)**.** *Let $(p,q)^{(1:T)}$ be a play history between a learner* $\mathbf{L}$ *with actions $\mathcal{H}$ and a adversary* $\mathbf{A}$ *with actions $\mathcal{G}$. Suppose $T\varepsilon \geq WReg_{\mathbf{L}}\left((p,q)^{(1:T)}\right)$ and $\sum_{t=1}^{T}\mathcal{L}^{(t)}(p^{(t)}) + T\varepsilon \geq \sum_{t=1}^{T}\max_{\ell^*}\mathcal{L}^*(p^{(t)})$. Then there exists a $t\in[T]$ where $p^{(t)}$ is a $2\varepsilon$-optimal solution for the multi-objective learning problem $(\mathcal{G},\mathcal{H})$.*

*Proof.* We will show that there is a timestep $t\in[T]$ where the no-regret learner's strategy is nearly min-max optimal $\max_{\ell^*\in\mathcal{G}}\mathcal{L}^*(p^{(t)}) \leq \min_{h^*\in\mathcal{H}}\max_{\ell^*\in\mathcal{G}}\mathcal{L}^*(h^*)+\varepsilon'$ using a proof by contradiction. Therefore assume to the contrary that for all $t\in[T]$ we have $\max_{\ell^*\in\mathcal{G}}\mathcal{L}^*(h^{(t)}) > \min_{h^*\in\mathcal{H}}\max_{\ell^*}\ell^*(h^*)+\varepsilon'$. It follows that,

$$\begin{aligned} T\varepsilon &> \sum_{t=1}^{T}\mathcal{L}^{(t)}(p^{(t)}) - \min_{h^*\in\mathcal{H}}\max_{\mathcal{L}^*} T\mathcal{L}^*(h^*) \\ &> \sum_{t=1}^{T} c\cdot\min_{h^*\in\mathcal{H}}\max_{\ell^*\in\mathcal{G}}\mathcal{L}^*(h^*) - \varepsilon + c\varepsilon' - T\min_{h^*\in\mathcal{H}}\max_{\mathcal{L}^*}\mathcal{L}^*(h^*), \end{aligned}$$

where the first inequality is our assumption by contradiction and the second inequality is because we have both the adversary best-responding and the learner no-regretting. This yields the contradiction

$$T(2\varepsilon - c\varepsilon') > Tc\cdot\mathrm{OPT} - \min_{h^*\in\mathcal{H}}\max_{\ell^*} T\ell^*(h^*) \to T(2\varepsilon + (1-c)\mathrm{OPT} - c\varepsilon') = 0 > 0.$$

□

**Remark 3.3.** *An analogy of Lemma B.2 also holds if the best-responding player is replaced by a weak learning oracle. In this case, assuming to the contrary that*

$$\sum_{t=1}^{T}\mathcal{L}^{(t)}(h^{(t)}) > \sum_{t=1}^{T}\min_{h^*}\max_{\ell^*}\mathcal{L}^*(h^*)+\varepsilon,$$

*immediately gives the contradiction,*

$$T\varepsilon \geq \sum_{t=1}^{T}\mathcal{L}^{(t)}(h^{(t)}) - \min_{h^*\in\mathcal{H}}\max_{\ell^*} T\mathcal{L}^*(h^*) > T\varepsilon.$$

Since Lemma 3.2 shows that there is some timestep where the no-regret learner plays a close-to-optimal strategy, we note that testing for which timestep this occurs at involves trivial sample complexity.

**Lemma 3.4.** *Take any set of hypotheses $p^{(1:T)}$ where at least one choice of $p^{(t)}$ is $\varepsilon$-optimal. The following procedure* $\textsc{Find}(\mathcal{H},\mathcal{G})$ *identifies a $2\varepsilon$-optimal solution with probability at least $1-\delta$, and with only $N = O\left(\varepsilon^{-2}\ln(T\left|\mathcal{G}\right|/\delta)\right)$ samples: return $p^{(t^*)}$ where $t^* = \arg\min_{t\in[T]}\max_{\ell\in\mathcal{G}}\widehat{\mathcal{L}}_S(p^{(t)})$ and $S$ is a set of $N$ samples drawn i.i.d. from $D$. This can also be implemented with $T$ calls to an agnostic learning oracle over $\mathcal{G}$ and 1 call to an agnostic learning oracle over $[T]$.*

**Designing the dynamics.** As we show in Section 4, existing multi-calibration algorithms, and our proposed algorithms, all can be described in terms of one of the above game dynamics. How these dynamics are implemented largely depends on what online learning and best-response guarantees are possible for each player. There must always be a player that leads and a player that follows, with the need to be no-regret being easier in general for the follower to meet. Indeed, in the calibration problems we will study, the learner readily attains zero-regret if allowed to follow, motivating us to have the adversary move first in most algorithms. However, once we turn to finding deterministic calibrated predictors as afforded by Lemma 3.2, it is necessary for the adversary to best-respond. This requires the learner to lead. In Lemma 4.2, we give a no-weak-regret learning algorithm that works for any calibration-like loss function.

When computing the adversary's best response, we often directly call agnostic learning oracles. When we can directly access the adversary's loss set $\mathcal{G}$, we use a standard tool from adaptive data analysis to implement to compute adversary best-response over an adaptive sequence of size $T$, with a small number of samples.

**Lemma 3.5** (Adaptive Data Analysis [1])**.** *There is an algorithm that for any distribution $D$, any set of loss functions $\mathcal{G} : \mathcal{H} \times \mathcal{X} \to [0,1]$, and any adaptive sequence of hypotheses $h^{(1:T)}$, with probability at least $1-\delta$ returns $\ell^{(1:T)}$ such that for all $t \in [T]$, $\mathcal{L}^{(t)}(h^{(t)}) \geq \max_{\ell^* \in \mathcal{G}} \mathcal{L}^*(h^{(t)}, z) - \varepsilon$, using no more samples from $D$ than:*

$$O\left(\frac{\sqrt{T}}{\varepsilon^2}\log\left(\frac{|\mathcal{G}|}{\varepsilon}\right)\log^{3/2}\left(\frac{1}{\varepsilon\delta}\right)\right) \approx \widetilde{O}\left(\frac{\sqrt{T}\ln(|\mathcal{G}|/\delta)}{\varepsilon^2}\right).$$

# 4 Calibration Guarantees

In this section, we use game dynamics to derive accessible proofs of, and more efficient algorithms for, multi-calibration and moment calibration. Both belong to an especially tractable class of *linear multi-objective learning problems* in which losses take the form $\ell(h,(x,y)) = \langle f_\ell(h,x), h(x) - g(y)\rangle$, where $f_\ell : \mathcal{H} \times \mathcal{X} \to [-1,1]^k$ and $g : \mathcal{Y} \to \Delta_k$.

**Best-responding to linear objectives.** A classical observation about these problems—dating back to Blackwell and Foster and formalized recently by Hart [17]—is that learners faced with linear objectives selected by an adversary can compute a best-response without any samples. This is a direct consequence of the minimax theorem.

**Fact 4.1** (Hart [17])**.** *Consider any online learning problem where an adversary chooses $\ell(h,(x,y)) = \langle f_\ell(h,x), h(x) - g(y)\rangle$ from a set of linear losses $\mathcal{G}$ and nature chooses datapoints $(x,y)$. Then for any mixed strategy of the adversary $q \in \Delta\mathcal{G}$ and any $r \in \mathbb{N}$, there is an explicitly constructable mapping $B_r : \Delta\mathcal{G} \to (\Delta\mathcal{Y})^{\mathcal{X}}$ that provides a (near-)zero regret strategy for the learner against any datapoint that nature may draw: $\forall z \in \mathcal{Z} : q(B_r(q), z) \leq 1/r$. Moreover, $B_r$ is distribution-free and can be computed without samples.*

It has immediate implications for finding non-deterministic solutions, since we can now thus implement no-regret vs no-regret dynamics with the learner following and have no sample complexity dependence on the learner (e.g., on $\mathcal{X}$ or $\mathcal{H}$).

**No-regret learning on linear objectives.** We now strengthen the previous result by proving that linear objectives also imply an online learning strategy for learners that does not require

randomization, can be computed without samples, and gives a weak-regret bound of order $O\left(\sqrt{T}\right)$ and *independent of domain complexity (e.g., $|\mathcal{X}|$).* This will allow us to implement game dynamics where the learner leads and the adversary follows.

**Lemma 4.2.** *Consider any linear multi-objective learning problem $(\mathcal{G}, \mathcal{H})$ where objectives are symmetric: for every $\ell \in \mathcal{G}$ and permutation $\sigma : [k] \to [k]$ there is a $\ell' \in \mathcal{G}$ where $f_\ell(\cdot) = \sigma(f_{\ell'}(\cdot))$. There is an online learning algorithm that only plays deterministic strategies and, against any adaptive adversary choosing losses $\ell^{(t)}$, guarantees the weak regret bound $WReg \le O(\sqrt{\ln(k)T})$.*

*Proof.* We will construct a predictor $h^{(t)} : \mathcal{X} \to \Delta_k$ by describing its behavior on each point $x$ in the domain. We interpret $h^{(t)}(x)$ as a distribution over $k$ actions returned by a no-regret algorithm on an adaptive sequence of losses, $\gamma_x^{(1)}, \ldots, \gamma_x^{(t-1)}$, where $\gamma_x^{(\tau)} : [k] \to [-1, 1]$. In particular, we consider the loss sequence $\gamma_x^{(\tau)}(\widehat{y}) \coloneqq \left\langle f^{(t)}(h^{(t)}, x), \widehat{y} \right\rangle$ for all $\widehat{y} \in [k]$. We note that $\gamma_x^{(\tau)}$ is a valid loss function for adaptive adversaries, who can observe the player's mixed strategy $h^{(t)}(x)$ for all $x$, but not the realized action $\widehat{y}$. Using Hedge, or any similar no-regret algorithm for $k$ actions, we have that $\sum_{t=1}^{T} \mathbb{E}_{\widehat{y}^{(t)} \sim h^{(t)}} \left[\gamma_x^{(t)}(\widehat{y}^{(t)})\right] \le \min_{y^* \in [k]} \sum_{t=1}^{T} \gamma_x^{(t)}(y^*) + O(\sqrt{\ln(k)T})$. Writing $\gamma_x^{(t)}$'s definition and using its linearity to substitute $\widehat{y}^{(t)} \sim h^{(t)}$ for $h^{(t)}$, we have,

$$O(\sqrt{\ln(k)T}) \ge \sum_{t=1}^{T} \left\langle f^{(t)}(h^{(t)}, x), h^{(t)} \right\rangle - \min_{y^* \in [k]} \sum_{t=1}^{T} \left\langle f^{(t)}(h^{(t)}, x), y^* \right\rangle.$$

Taking an expectation over $D$, and subtracting the true labels $g(y)$ from each side, we obtain

$$O\left(\sqrt{\ln(k)T}\right) \ge \mathbb{E}\left[\sum_{t=1}^{T} \left\langle f^{(t)}(h^{(t)}, x), h^{(t)} - g(y) \right\rangle\right] - \underbrace{\min_{h^*} \mathbb{E}\left[\sum_{t=1}^{T} \left\langle f^{(t)}(h^{(t)}, x), h^*(x) - g(y) \right\rangle\right]}_{(a)}.$$

We note that $(a) \le 0$ since $g(y)$ is conditionally (on $x$) independent of $f^{(t)}(h^{(t)}, x)$ so the minimizer can always choose $h^*(x)$ to be the Bayes classifier $\mathbb{E}[g(y) \mid x]$. This leaves to be shown that $(a)$ is less than OPT by proving $\mathrm{OPT} \coloneqq \min_{h^*} \max_{\ell^* \in \mathcal{G}} \mathbb{E}\left[\langle f^*(h^*, x), h^*(x) - g(y)\rangle\right]$ is nonnegative. We will do so by the symmetry of $\mathcal{G}$. Let $h^*$ and $\ell^*, f^*$ correspond to the values attaining OPT. We have,

$$\begin{aligned}
&\sum_{\sigma \in [k] \to [k]} \mathbb{E}\left[\langle \sigma(f^*(h^*, x)), h^*(x) - g(y)\rangle\right] \\
&= \mathbb{E}\left[\sum_{\sigma \in [k] \to [k]} \sum_{j \in [k]} f^*(h^*, x)_{\sigma(j)} (h^*(x)_j - g(y)_j)\right] \\
&= \mathbb{E}\left[\sum_{j \in [k]} \left(\sum_{\sigma \in [k] \to [k]} f^*(h^*, x)_{\sigma(1)}\right) (h^*(x)_j - g(y)_j)\right] \\
&= \mathbb{E}\left[\left(\sum_{\sigma \in [k] \to [k]} f^*(h^*, x)_{\sigma(1)}\right) \sum_{j \in [k]} (h^*(x)_j - g(y)_j)\right] \\
&= 0,
\end{aligned}$$

where the second transition is because, for any vector $v$, $\sum_\sigma v_{\sigma(j)}$ is independent of $j$; and the last transition is because $h^*$ and $g(y)$ are both distributions over $[k]$ and each sum to one. This equation

implies there is a permutation $\sigma$ such that $\mathbb{E}\left[\langle \sigma(f^*)(h^*, x), h^*(x) - g(y)\rangle\right] \geq 0$. Since $\sigma(f^*) \in \mathcal{G}$ by symmetry, it must follow that $\mathrm{OPT} \geq 0$. □

### 4.1 Multi-calibration

Multi-calibration is an example of a linear multi-objective learning problem, one that extends calibration to ask that one's predictor be calibrated on multiple subsets of the domain [18]. We will use the following definition of multi-calibration.[1] To simplify the notation of level sets, we will use $V_\lambda \coloneqq [0, \lambda, \ldots, \lambda \lceil 1/\lambda \rceil]$ to denote the $\lambda$-discretization of $[0, 1]$.

**Definition 4.3.** *Fix some $\varepsilon, \lambda, k > 0$ and $\mathcal{S} \subseteq \{0,1\}^{\mathcal{X}}$. Suppose $\mathcal{Y} = [0,1]^k$. A $k$-class predictor $h : \mathcal{X} \to \mathcal{Y}$ is* $(\mathcal{S}, \varepsilon, \lambda)$-multi-calibrated *if for all $S \in \mathcal{S}$, $v \in V_\lambda^k$, and $j \in [k]$ we have $\left|\mathbb{E}_{(x,y)\sim D}\left[(h(x) - y)_j \cdot 1[h(x) \in v, x \in S]\right]\right| \leq \varepsilon$. We refer to the special case of $k = 2$, i.e., $\mathcal{Y} = [0,1]$, simply as* multi-calibration.

**Fact 4.4.** *Let $(\mathcal{S}, \varepsilon, \lambda)$ be a $k$-class multi-calibration problem on distribution $D$. $h \in \Delta_k^{\mathcal{X}}$ is a $(\mathcal{S}, \varepsilon, \lambda)$-multi-calibrated predictor if $h$ is $\varepsilon$-optimal for the multi-objective problem $(\mathcal{G}, \Delta_k^{\mathcal{X}})$ where $\mathcal{G} \coloneqq \left\{(h(x)_i - y(x)_i) \cdot 1[h(x) \in v, x \in S] \mid \forall S \in \mathcal{S}, v \in V_\lambda^k, i \in [k]\right\}$.*

This means we can use no-regret dynamics (Lemma 3.1) and the best-response method for linear learners (Fact 4.1) to immediately provide a $\widetilde{O}\left(\ln(|\mathcal{S}|)/\varepsilon^2\right)$ sample complexity bound.

**Corollary 4.5.** *For any $k$-class multi-calibration problem, the following algorithm takes no more than $O\left(\varepsilon^{-2}\left(\ln(|\mathcal{S}|/\delta) + k\ln(1/\lambda)\right)\right)$ samples from $D$ and returns a nondeterministic predictor $p \in \Delta(\mathcal{X} \to \Delta_k)$ such that w.p. at least $1 - \delta$, $p$ is a $(\mathcal{S}, \varepsilon, \lambda)$-multi-calibrated predictor:*

No-Regret vs No-Regret

*Construct the corresponding multi-objective learning problem $\mathcal{G}$ as in Fact 4.4. For $T = O\left(\varepsilon^{-2}\left(\ln(|\mathcal{S}|/\delta) + k\ln(1/\lambda)\right)\right)$ rounds, use a no-regret algorithm of regret $\sqrt{T\ln(|\mathcal{G}|)}$ to choose $\ell^{(t)}$, use the best-response $B_r(\ell^{(t)})$ of Fact 4.1 to set $h^{(t)}$, and sample $z^{(t)} \sim D$ to provide a feedback vector to the adversary. Let $p$ be uniform over $h^{(1:T)}$.*

*Proof.* As this result is immediate from Lemma 2.1, Fact 4.4, and Lemma 3.1, we will instead review how the claimed algorithm maps into these lemmas as an example of how to construct these types of game dynamics.

First, we want to verify that the multi-objective learning problem posed by multi-calibration is linear and has objectives taking the form $\ell(h, (x, y)) = \langle f_\ell(h, x), h(x) - g(y)\rangle$. As constructed by Fact 4.4, this is indeed the case with each objective $\ell \in \mathcal{G}$ corresponding to choosing $g$ as the identity function, having $\mathcal{G} \cong \mathcal{S} \times V_\lambda^k \times k$, and letting $f$ be an indicator function. This means that the best response $B_r(x)$ indeed exists for the learner, and we can choose a sufficiently large $r$ so that the learner has distributional regret of effectively 0: $\mathrm{Reg}_{\mathbf{L}}((h, \ell)^{(1:T)}) \approx 0$. Similarly, we know that the adversary has empirical regret of $O(\sqrt{T\ln(|\mathcal{G}|)})$ by construction; Hedge would suffice for this, for example. Recall that at each iteration we sample a single datapoint to estimate the adversary's payoff vector; by linearity, this estimate is naturally unbiased. By Lemma 2.1, because the adversary is running an online learning algorithm, its generalization error is of the same order as its regret; that is, the adversary has distributional regret $O(\sqrt{T\ln(|\mathcal{G}|)})$. Since the learner and adversary

[1] We use the $(\alpha, \lambda)$-definition of multi-calibration [18] and will always state our results and other work's results for this $(\alpha, \lambda)$ definition. Also, we use uniformly spaced calibration bins throughout, but our results directly apply to uniformly occupied bins as well.

have distributional regret bounded by $O(\sqrt{T \ln(|\mathcal{G}|)})$, Lemma 3.1 confirms the iteration complexity and—since we only sample one datapoint per iteration—sample complexity of our dynamics. □

Noarov et al. [25], Gupta et al. [15] achieve a matching rate for the binary setting ($k = 2$) by using a more involved argument about minimizing an exponential potential function over $\mathcal{G}$. Upon closer inspection, their potential argument is a reiteration of Hedge's regret bound. Corollary 4.5 establishes a more general and direct approach for obtaining these guarantees. An important caveat of this approach generally is that it yields predictors that are nondeterministic. However, we observe that one can recover a deterministic predictor from these procedures by using the discreteness of the calibration bins.

**Proposition 4.6.** *For any $k$-class multi-calibration problem, the following algorithm takes no more than $O\left(\lambda^{-2}\varepsilon^{-2}\left(\ln(|\mathcal{S}|/\delta) + k\ln(1/\lambda)\right)\right)$ samples from $D$ and returns a deterministic predictor $h^* : \mathcal{X} \to [0,1]^k$ such that w.p. at least $1-\delta$, $h^*$ is $(\mathcal{S}, \varepsilon, \lambda)$-multi-calibrated:*

No-Regret vs No-Regret with Majority Vote

*Let $h^{(1:T)}$ be the predictors returned by the algorithm of Corollary 4.5. Construct $h^*$ by defining $h^*(x) \coloneqq \text{Mean}(\{h^{(t)}(x) \mid t \in [T], h^{(t)}(x) \in [v^*(x), v^*(x)+\lambda)\})$ where $v^*(x) \coloneqq \text{Mode}\left(\left\{\lfloor h^{(t)}(x)/\lambda \rfloor\right\}_{t\in[T]}\right)$ is the majority vote on which bin $h^*(x)$ should lie in.*

*Proof.* The linear multi-objective learning expression of multi-calibration has the additional property that, for every $\ell \in \mathcal{G}$, $f_\ell$ only depends on $h$ locally: $\ell(h,(x,y)) = \langle f_\ell(\mathbf{h}(\mathbf{x}), x), h(x) - y\rangle$. For each datapoint $x \in \mathcal{X}$, let $v^*(x)$ be the majority vote about which bin the prediction for $x$ should fall in: $v^*(x) \coloneqq \text{Maj}\left(\left\{\lfloor h^{(t)}(x)/\lambda \rfloor\right\}_{t\in[T]}\right)$. Then, construct predictor $h^* : \mathcal{X} \to \mathcal{Y}$ by setting $h^*(x) \coloneqq \frac{1}{|\mathcal{I}(x)|}\sum_{t\in\mathcal{I}(x)} h^{(t)}(x)$ where $\mathcal{I}(x) \coloneqq \{t \in [T] \mid h^{(t)}(x) \in [v^*(x), v^*(x)+\lambda)\}$. We will see that this only leads to a $1/\lambda$ degradation in our risk bound $\varepsilon$. For any objective $\ell \in \mathcal{G}$, we can exploit the fact that $\ell$ is linear in $h(x)$ for all values of $h(x)$ belonging to the same bin:

$$
\begin{aligned}
T\varepsilon &\geq \sum_{t=1}^{T} \mathop{\mathbb{E}}_{x}\left[\ell(h^{(t)}(x), x)\right] \\
&= \mathop{\mathbb{E}}_{x}\left[\sum_{t\in\mathcal{I}(x)} \ell(h^{(t)}(x), x)\right] \\
&= \mathop{\mathbb{E}}_{x}\left[|\mathcal{I}(x)|\,\ell\left(\frac{1}{|\mathcal{I}(x)|}\sum_{t\in\mathcal{I}(x)} h^{(t)}(x), x\right)\right] \\
&= \mathop{\mathbb{E}}_{x}\left[|\mathcal{I}(x)|\,\ell\left(\sum_{t\in\mathcal{I}(x)} h^{(t)}(x), x\right)\right] \\
&\geq \left\lceil\frac{T}{\lambda}\right\rceil \mathop{\mathbb{E}}_{x}\left[\sum_{x\in\mathcal{X}} |\mathcal{I}(x)|\,\ell(v^*(x), x)\right].
\end{aligned}
$$

The claim then follows by recalling that the uniform distribution on $h^{(1:T)}$ provided by Corollary 4.5 is already $\varepsilon$-optimal. □

To learn deterministic multi-calibrated predictors with even fewer samples, we turn to No-Regret vs Best-Response dynamics. Whereas previously we had the adversary lead in the dynamics used for Corollary 4.5, in order to use NRBR dynamics, the learner must lead. In order to recover a deterministic predictor, the learner must not only lead, but also achieve vanishing regret while only playing deterministic strategies. In Lemma 4.2, we showed that such an online learning strategy indeed exists for linear multi-objective learning problems like calibration. In the following corollary, we put these pieces together, using an agnostic learning oracle to implement the best-response strategy of the adversary.

**Corollary 4.7.** *For any $k$-class multi-calibration problem, the following algorithm makes $O\left(\varepsilon^{-2}\ln(k)\right)$ calls to an agnostic learning oracle and returns a deterministic $k$-class predictor $h^* : \mathcal{X} \to [0,1]^k$ such that w.p. at least $1-\delta$, $h^*$ is $(\mathcal{S},\varepsilon,\lambda)$-multi-calibrated:*

Best-Response vs No-Regret with FIND

*Construct the corresponding multi-objective learning problem $\mathcal{G}$ as in Fact 4.4. For $T = O\left(\varepsilon^{-2}\ln(k)\right)$ rounds, use the no-regret algorithm of Lemma 4.2 to choose $h^{(t)}$ and call the agnostic learning oracle to choose $\ell^{(t)}$. Let $h^* = \text{FIND}(\{h^{(t)}\}_{t\in[T]}, \mathcal{G})$.*

*Proof.* As we showed in Corollary 4.5, $\mathcal{G}$ is a linear multi-objective expression of our multi-calibration problem. This means that Lemma 4.2 implies the existence of an online learning strategy for the learner with weak regret bound $O(\sqrt{\ln(k)T})$ and no use of randomization. Similarly, we know the adversary to be best-responding by assumption, because we delegate it's best-responses to an agnostic learning oracle. By Lemma 3.2, it follows that w.p. $1-\delta/2$ there exists a $t \in [T]$ where $h^{(t)}$ is $\varepsilon/2$-optimal. By Lemma 3.4, the procedure Find w.p. $1-\delta/2$ identifies a $h^{(t')}$ that is $\varepsilon$-optimal; moreover, this $h^{(t')}$ is a deterministic predictor. The sample complexity claim follows since we only make one oracle call per iteration and Find can be implemented with either only $\ln(T/\delta)/\varepsilon^2$ samples or $T$ more oracle calls. □

Corollary 4.7 easily extends to variants of multi-calibration like degree-$k$ multi-calibration [14]. Note that one can translate the oracle complexity guarantee of Corollary 4.7 into a sample complexity by implementing the best-responding adversary with an adaptive data analysis algorithm.

**Corollary 4.8.** *Using Lemma 3.5 to implement an agnostic learning oracle, Corollary 4.7's algorithm takes $\widetilde{O}(\varepsilon^{-3}(\sqrt{\ln(k)}\ln(1/\delta)(\ln(|\mathcal{S}|)+k\ln(1/\lambda))))$ samples and returns a deterministic $k$-class predictor $h^* : \mathcal{X} \to [0,1]^k$ where, w.p. $1-\delta$, $h^*$ is $(\mathcal{S},\varepsilon,\lambda)$-multi-calibrated.*

*Proof.* This claim follows immediately from Corollary 4.7 and Lemma 3.5. □

These corollaries improve upon Gopalan et al. [14]'s oracle complexity upper bound of $O\left(k/\varepsilon^2\right)$ with a bound of $O\left(\ln(k)/\varepsilon^2\right)$, an exponential reduction in the dependence on the number of classes $k$. In concurrent work, Dwork et al. [7] also showed that $O\left(\ln(k)/\varepsilon^2\right)$ oracle calls are sufficient for multi-class multi-calibration, with an algorithm similar to Corollary 4.7 but using Dwork et al. [6]'s outcome indistinguishability and working with specific multicalibration objectives. Our result covers a more general class of problems, as we will see in the next section on moment calibration.

Together, Corollaries 4.5 and 4.7 each reflect previously distinct philosophies to multi-calibration algorithm design: using online learning to quickly find non-deterministic predictors [15, 25], and using potential-based arguments to more slowly find deterministic predictors [18, 6]. Our result shows that these two approaches are closely related game dynamics.

### 4.2 Moment calibration

We now extend our study of game dynamics from two-player games to games with sets of learners and adversaries. This allows us to address a broader class of multi-objective problems while leveraging the same helpful game dynamics as before (easily extending their proofs).

At a high level, this setting involves decomposing a multi-objective learning problem $(\mathcal{D}, \mathcal{G}, \mathcal{H})$ into $b$ components $\mathcal{H} \coloneqq \prod_{a\in[b]} \mathcal{H}_a$ and $\mathcal{G} \coloneqq \prod_{a\in[b]} \mathcal{G}_a$, and introducing a pair of learner and adversary players for each of the $b$ components. Each pair of learner and adversary, who have agency over their own choices of actions, has a loss determined by not only their actions but also the fixed actions of other players. We capture this as a single multi-objective learning problem with a history of play $\{h^{(t)}_{1:b}, \ell^{(t)}_{1:b}\}_{t\in[T]}$ such that for any $a \in [b]$, learner $a$'s actions $h^{(t)}_a : \mathcal{X} \to [0,1]^k$ belong to $\mathcal{H}_a$ and the corresponding adversary $\mathbf{A}_a$'s actions $\ell^{(t)}_a : \mathcal{H}_a \times \mathcal{X} \times \mathcal{Y} \to [0,1]$ belong to $\mathcal{G}_a$. To emphasize that the loss of any learner $a$ still depends on the actions of other learners, we use the notation $\ell_a(h_a, h_{-a}, z)$—or when clear from the context—the notation $\ell_a(h, z)$ for $z \in \mathcal{X} \times \mathcal{Y}$. Given the decomposition of the above multi-objective learning problem across its components, we sometimes construct solutions that are optimal on a subset of the coordinates only, and combine these solutions later on.

**Definition 4.9.** *For a multi-objective learning problem $(\mathcal{G}, \mathcal{H})$ where $\mathcal{H} = \prod_{i\in[b]} \mathcal{H}_i$, a solution that is* component-wise *$\varepsilon$-optimal w.r.t. to a subset of its components $\mathcal{I} \subseteq [b]$ is a hypothesis $p$ satisfying, for every $i \in \mathcal{I}$, that $\mathcal{L}^*(p) - \min_{h^*_i \in \mathcal{H}_i} \mathcal{L}^*(h^*_i, p_{-i}) \le \varepsilon$.*

The *moment multi-calibration problem* introduced by Jung et al. [19] is concerned not only with having calibrated estimates of the mean of a population, but also of higher moments. This is motivated by confidence interval estimation, as moment estimates provide rough Chernoff bounds.[2]

**Definition 4.10.** *Fix some $\varepsilon, \lambda, r > 0$ and $\mathcal{S} \subseteq \{0,1\}^{\mathcal{X}}$. Suppose $\mathcal{Y} = [0,1]$. The pair of predictors $h_\mu : \mathcal{X} \to [0,1]$ and $h_m : \mathcal{X} \to [0,1]^r$ is* $(\mathcal{S}, \varepsilon, \lambda)$-mean-conditioned moment-multi-calibrated *if for all groups $S \in \mathcal{S}$, level sets $v_\mu, v_m \in V_\lambda$, and moments $a \in [r]$, we have that the mean estimator is calibrated $|\mathbb{E}\left[(h_\mu(x) - y) \cdot 1[h_\mu(x) \in v_\mu, h_{m,a}(x) \in v_m, x \in S]\right]| \le \varepsilon$ and the moment estimators are calibrated $|\mathbb{E}\left[(h_{m,a}(x) - (y - h_\mu(x))^a) \cdot 1[h_\mu(x) \in v_\mu, h_{m,a}(x) \in v_m, x \in S]\right]| \le \varepsilon$.*

In this definition, we use $\mathcal{Y} = [0,1]$ to denote the probability of predicting a label of 1. To match our notation of multi-objective learning for the remainder of this section we will instead consider $\mathcal{Y} = \Delta_2$ such that a prediction probability $y$ is represented as a vector $(y, 1-y) \in \Delta_2$. In this language, moment multi-calibration constitutes a linear multi-objective learning problem with $k = 2$ by convention; in the multi-component setting, linear objectives mean that, for every $a \in [r], \ell \in \mathcal{G}_a$, we can write $\ell_a(h, (x, y)) \coloneqq \langle f_a(h, x), h_a(x) - g(y)\rangle$.

A challenging aspect of moment calibration is that $h_\mu$'s optimality criterion depends on $h_m$, and vice versa. Jung et al. [19] handled this with a careful analysis of alternating gradient descent from a single-agent perspective, overall requiring a suboptimal $\widetilde{O}\left(r \ln(1/\delta)/\varepsilon^4\right)$ number of calls to an agnostic learning oracle. As we work with game dynamics, we can simply introduce $h_\mu$ and $h_m$ as different players to get, as a corollary of our previous results, a oracle complexity for deterministic moment calibration of only $\widetilde{O}\left(r \ln(1/\delta)/\varepsilon^2\right)$.

[2] We use Jung et al. [19]'s "pseudo" definition of moment calibration—where moments are defined w.r.t. estimated means—because all existing algorithms work by guaranteeing "pseudo" rather than literal calibration anyways. All results results are stated terms of Definition 4.10.

**Fact 4.11.** *Take a moment calibration problem $(\mathcal{S}, \lambda)$. For $S \in \mathcal{S}$, $v, w \in V_\lambda$, $a \in [r]$, and $i \in [2]$, include in $\mathcal{G}_\mu$ and $\mathcal{G}_m$, respectively, the objectives:*

$$\ell_{S,v,w,i,\mu}((h_\mu, h_m), x) \coloneqq 1[h_\mu(x) \in v]1[h_{m,a}(x) \in w]1[x \in S](h_\mu(x)_i - y_i)$$
$$\ell_{S,v,w,i,a}((h_\mu, h_m), x) \coloneqq 1[h_\mu(x) \in v]1[h_{m,a}(x) \in w]1[x \in S](h_{m,a}(x)_i - (y_1 - h_\mu(x)_1)^a).$$

*If $h$ is a component-wise $\varepsilon$-optimal solution for both learners $h_\mu, h_m$, then it is also $(\mathcal{S}, \varepsilon, \lambda)$-mean-conditioned moment multi-calibrated.*

**Corollary 4.12.** *For any $r$-moment calibration problem, the following algorithm takes $O\left(\varepsilon^{-2} r\right)$ calls to an agnostic learning oracle (analogously, $\widetilde{O}\left(\varepsilon^{-3}\sqrt{r}\ln(|\mathcal{S}|/\lambda)\ln(1/\delta)\right)$ number of samples) and returns a deterministic predictor $p : \mathcal{X} \to [0,1]^r$ such that w.p. at least $1-\delta$, $p$ is a $(\mathcal{S}, \varepsilon, \lambda)$-mean-conditioned moment multi-calibrated predictor:*

Best-Response vs No-Regret with FIND

*Construct the corresponding multi-objective learning instance as in Fact 4.11. For $T = O\left(r/\varepsilon^2\right)$ rounds, use agnostic learning oracles to choose a $\ell_m^{(t)} \in \mathcal{G}_m$ and $\ell_\mu^{(t)} \in \mathcal{G}_\mu$, use two copies of the no-regret algorithm of Lemma 4.2 to choose $h_m^{(t)} : \mathcal{X} \to (\Delta_2)^r$ and to choose $h_\mu^{(t)} : \mathcal{X} \to \Delta_2$, and sample a $z^{(t)} \sim D$ to provide a feedback vector to the adversary. Call procedure* FIND$(\{h_m^{(t)}, h_\mu^{(t)}\}_{t\in[T]}, \mathcal{G})$ *to find an $\varepsilon$-optimal choice of $h_m^*, h_\mu^*$, setting $h^* = [h_\mu^*, (h_m^*)_2, \ldots, (h_m^*)_r]$.*

*Proof.* In the claimed algorithm, we instantiate no-regret best-response dynamics with two sets of learners and adversaries ($b = 2$), which respectively handle mean and moment prediction. Similarly to multi-calibration, we see by Fact 4.11 that moment calibration is indeed expressible as a linear multi-objective learning problem where for each $\ell \in \mathcal{G}$, $f_\ell$ is a indicator functions, and $g$ is either the identity function (for $\mathcal{G}_\mu$) or $g(h, (x, y)) = (y - h_\mu(x))^a$ (for $\mathcal{G}_m$). This means that Lemma C.1 (note that we are using the generalization of Lemma 4.2; this is because $h_m$ is handling a multi-label problem) applies so we know each of the learners' strategies in these dynamics to guarantee them individual distributional regrets of $O\left(\sqrt{\ln(2^r)T}\right)$. Similarly, we know the adversaries to be best-responding by assumption, because we delegate their best-responses to their own agnostic learning oracles. By Lemma B.2, it follows that w.p. $1 - \delta/2$ there exists a $t \in [T]$ where $h_m^{(t)}$ and $h_\mu^{(t)}$ are both $\varepsilon/2$-optimal. By Lemma 3.4, the procedure Find w.p. $1 - \delta/2$ identifies a $h^{(t')}$ that is $\varepsilon$-optimal. □

As with multi-calibration, game dynamics allows us to unify online and batch treatments of moment calibration. By switching to no-regret vs no-regret dynamics, we next derive sample complexity rates matching Gupta et al. [15]'s rates of $O\left(\varepsilon^{-2}\ln(|\mathcal{S}|\, r/\lambda)\right)$ for finding nondeterministic predictors for moment calibration. Note that nondeterministic predictors for moments are less useful for building confidence intervals, so deterministic settings are likely more practical.

**Corollary 4.13.** *For any $r$-moment calibration problem, the following algorithm returns a nondeterministic predictor $p \in \Delta(\mathcal{X} \to [0,1]^r)$ and takes only $O\left(\varepsilon^{-2}\ln(|\mathcal{S}|\, r/\lambda\delta)\right)$ samples, such that w.p. $1-\delta$, $p$ is $(\mathcal{S}, \varepsilon, \lambda)$-mean-conditioned moment multi-calibrated:*

No-Regret vs No-Regret

*Construct the corresponding multi-objective learning instance as in Fact 4.11. For $T = O\left(\varepsilon^{-2}\ln(|\mathcal{S}|\, r/\lambda\delta)\right)$ rounds, use two no-regret algorithms of regret $\sqrt{T\ln(|\mathcal{G}|)}$ to choose a $\ell_m^{(t)} \in \mathcal{G}_m$*

*and $\ell_\mu^{(t)} \in \mathcal{G}_\mu$, use two copies of the no-regret algorithm of Lemma 4.2 to choose $h_m^{(t)} : \mathcal{X} \to (\Delta_2)^r$ and to choose $h_\mu^{(t)} : \mathcal{X} \to \Delta_2$, and sample a $z^{(t)} \sim D$ to provide a feedback vector to the adversary. Let $p(x)$ be the uniform distribution over $[(h_\mu)_1^{(t)}, (h_{m,2})_1^{(t)}, \dots (h_{m,r})_1^{(t)}]_{t\in[T]}$.*

*Proof.* This proof follows identically to Corollary 4.12. In this case, we instead use no-regret algorithms to implement our adversaries with low distributional regret, invoking Lemma A.1. □

# 5 New Considerations for Multi-Objective Learning

In this section, we show that embracing different types of game dynamics in multi-objective learning allows us to enrich the fairness guarantees afforded by multi-calibration and adjacent frameworks.

**Agnostic multi-calibration.** Multi-calibration defines a population/group as a set of included features $S \subseteq \mathcal{X}$. These definitions do not address common use cases where two individuals which belong to different identity groups share the same features, e.g., due to regulatory or practical reasons. To overcome this, we consider a multi-calibration setting that allows people with the same feature vectors to belong to different populations. Multi-calibration may not be possible in these cases since identity groups may have conflicting Bayes classifiers. Instead, we can relax to ask that one provide the best-possible calibration guarantee that can hold uniformly across all subpopulations.

**Definition 5.1.** *Fix some $\varepsilon, \lambda, k, u > 0$. Suppose $\mathcal{Y} = [0,1]^k$ and $\mathcal{X} = \mathcal{X}' \times \{0,1\}^u$. A $k$-class predictor $h : \mathcal{X}' \to \mathcal{Y}$ is* $(\varepsilon, \lambda)$-agnostic multi-calibrated *if $\mathcal{L}^*(h) \le \min_{h^*} \mathcal{L}^*(h^*) + \varepsilon$ where $\mathcal{L}^*(h) \coloneqq \max_{i\in[u], v\in V_\lambda^k, j\in[k]} \left|\mathbb{E}_{((x,w),y)} \left[(h(x) - y)_j \cdot 1[h(x) \in v, i \in w]\right]\right|$. Here, we limit ourselves to a hypothesis class $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}'}$ where*

The absence of the concept of a Bayes classifier in this setting rules out existing proofs of multi-calibration using potential arguments and outcome indistinguishability [see, e.g., 18, 6]. However, an advantage of building our technical approach on game dynamics is that our results immediately extend to provide agnostic multi-calibration guarantees. In fact, our algorithms and analysis for multi-calibration (namely Corollaries 4.7 and 4.5) extend directly to the non-agnostic case with comparable sample complexity. Note, however, that our linear objectives take the slightly different form $\ell(h, (x, w)) = \langle f_\ell(h, x), h(x) - g(y)\rangle$.

For non-deterministic predictors, agnostic multi-calibration is achievable with no change in our sample complexity bound. This is because Lemma 3.1 and Fact 4.1 carry through as usual.

**Corollary 5.2.** *For any $k$-class multi-calibration problem the algorithm of Corollary 4.5 takes $O\left(\varepsilon^{-2}\left(\ln(u/\delta) + k\ln(1/\lambda)\right)\right)$ samples from $D$ and returns a non-deterministic predictor $p \in \Delta(\mathcal{X} \to \Delta_k)$ such that w.p. $1-\delta$, $p$ is a $(\varepsilon, \lambda)$-agnostic-multi-calibrated predictor.*

For deterministic predictors, agnostic multicalibration requires more careful consideration. This is because the optimal deterministic agnostic multi-calibrated predictor can involve unnecessary, and sometimes undesirable, optimizations. Specifically, in the below remark, we construct an example where a learner is incentivized to lump points together into the same level set.

**Remark 5.3.** *Consider agnostic calibration on the learning problem where the label space is binary $k = 2$, there are two groups we want to be calibrated over $u = 2$, and the domain consists of three points $\mathcal{X} = \{x_1, x_2, x_3\}$ over which we define a uniform distribution. The two groups are both*

*supported on, and have the same label distribution on, point* $x_1$*:* $\Pr(y = 1 \mid x = x_1, w = 1) = \Pr(y = 1 \mid x = x_1, w = 2) = 0.1$*. The two groups are also supported on, but have opposing label distributions on, points* $x_2, x_3$*:* $\Pr(y = 1 \mid x = x_2, w = 1) = \Pr(y = 1 \mid x = x_3, w = 2) = 1$ *and* $\Pr(y = 2 \mid x = x_3, w = 1) = \Pr(y = 2 \mid x = x_2, w = 2) = 1$*. It seems like an optimal predictor should look like* $\Pr(\widehat{y} = 1 \mid x = x_1) = 0.1$ *and* $\Pr(\widehat{y} = 1 \mid x = x_2) = \Pr(\widehat{y} = 1 \mid x = x_3) = 0.5$*. However, the optimal agnostic multicalibrated predictor actually has* $|\Pr(\widehat{y} = 1 \mid x = x_1) - \Pr(\widehat{y} = 1 \mid x = x_2)| \leq \lambda$ *and vise versa for* $x_1$ *and* $x_3$*. This is because the learner is incentivized to ensure that its predictions for point* $x_1$ *and points* $x_2, x_3$ *all fall into the same calibration bin so as to dilute the non-realizability of points* $x_2, x_3$*.*

Thus, for deterministic predictors, we will relax from seeking $\varepsilon$-optimality. Rather, we will seek predictors that are near $\varepsilon$-optimal but allowed some slack if, within the predictor's calibration bins, there is a significant difference between the expected labels of each group. We will formalize this slack with the quantity $\Delta \coloneqq \max_{v \in V_\lambda^k} \max_{j \in [u]} \mathbb{E}_x \left[\mathrm{Cov}(1[h^*(x) \in v, j \in w], y \mid x)\right]$, describing the expected correlation between the event that a datapoint falls in calibration bin $v$ and belongs to group $j$, and the event that a label $y$ is returned. In our previous remark's example, this slack is sufficient to restore the optimal predictor as being $\Pr(\widehat{y} = 1 \mid x = x_1) = 0.1$ and $\Pr(\widehat{y} = 1 \mid x = x_2) = \Pr(\widehat{y} = 1 \mid x = x_3) = 0.5$.

**Corollary 5.4.** *For any k-class multi-calibration problem, the algorithm of Corollary 4.7 (with a modified Find subroutine) takes* $O\left(\varepsilon^{-2} \ln(k/\delta)\right)$ *calls to an agnostic learning oracle and returns a deterministic k-class predictor* $h^* : \mathcal{X} \to [0,1]^k$ *such that w.p.* $1 - \delta$*,* $h^*$ *is* $(\varepsilon - \Delta, \lambda)$*-agnostic-multi-calibrated, where* $\Delta \coloneqq \max_{v \in V_\lambda^k} \max_{j \in [u]} \mathbb{E}_x \left[\mathit{Cov}(1[h^*(x) \in v, j \in w], y \mid x)\right]$*.*

**Group-conditional multi-calibration.** Another subtle aspect of multi-calibration definitions is that the strength of the guarantees is proportional to the size of a subgroup in a prediction bracket (note the use of the indicator function in Definition 4.3). This can lead to predictors that are uncalibrated on underrepresented/rare groups, and motivates our definition of *conditional* multi-calibration below.

**Definition 5.5.** *Fix some* $\varepsilon, \lambda, k > 0$*. Suppose* $\mathcal{Y} = [0,1]^k$ *and* $\mathcal{S} \subseteq \{0,1\}^{\mathcal{X}}$*. A k-class predictor* $h : \mathcal{X} \to \mathcal{Y}$ *is* $(\varepsilon, \lambda)$-conditional multi-calibrated *if for all* $S \in \mathcal{S}$*,* $v \in V_\lambda^k$*, and* $j \in [k]$*, we have* $\left|\mathbb{E}_{(x,y)\sim D}\left[(h(x) - y)_j \cdot 1[h(x) \in v] \mid x \in S\right]\right| \leq \varepsilon$*.*

Since this stronger notion of multi-calibration necessarily requires that algorithms be able to sample from the data distribution conditional on the events $x \in S$, we will assume access to example oracles $\mathrm{EX}_S$ for each $S \in \mathcal{S}$ that each yield samples from the conditional distribution $D \mid x \in S$. Accordingly, we will reduce conditional multi-calibration to a multi-distribution multi-objective learning problem of form $(\mathcal{D}, \mathcal{G}, \mathcal{H})$ rather than the single-distribution problems $(\mathcal{G}, \mathcal{H})$ we have considered previously. Here, $\mathcal{D}$ refers to a set of data distributions for which example oracles are available and we define the multi-distribution multi-objective loss as $\mathcal{L}^*(h) \coloneqq \max_{D \in \mathcal{D}} \max_{\ell \in \mathcal{G}} \ell_{D,\ell}(h)$.

**Fact 5.6.** *Let* $(\mathcal{S}, \varepsilon, \lambda)$ *be a k-class multi-calibration problem on distribution* $D$*.* $h \in \Delta_k^{\mathcal{X}}$ *is a* $(\mathcal{S}, \varepsilon, \lambda)$*-conditional multi-calibrated predictor if* $h$ *is* $\varepsilon$*-optimal for the multi-objective problem* $\left(\{D \mid x \in S\}_{S \in \mathcal{S}}, \mathcal{G}, \Delta_k^{\mathcal{X}}\right)$ *where* $\mathcal{G} \coloneqq \left\{(h(x) - y(x))_j \cdot 1[h(x) \in v] \mid v \in V_\lambda^k, x \in S, j \in [k]\right\}$*.*

For non-deterministic multi-calibration, this reduction implies, as one would expect, a linear (in $|\mathcal{S}|$) increase in sample complexity upper bounds.

**Corollary 5.7.** *For any k-class multi-calibration problem* $(\mathcal{S}, \varepsilon, \lambda)$ *the following algorithm makes* $O\left(\varepsilon^{-2} |\mathcal{S}| \left(\ln(|\mathcal{S}|/\delta) + k \ln(1/\lambda)\right)\right)$ *calls to example oracles in* $\{EX_S\}_{S \in \mathcal{S}}$ *and returns a non-deterministic*

*predictor $p \in \Delta(\mathcal{X} \to \Delta_k)$ such that w.p. $1-\delta$, $p$ is a $(\mathcal{S}, \varepsilon, \lambda)$-conditionally multi-calibrated predictor:*

No-Regret vs No-Regret

*Construct the multi-objective problem $(\mathcal{D}, \mathcal{G}, \mathcal{H})$ of Fact 5.6. For $T = O\left(\varepsilon^{-2}\left(\ln(|\mathcal{S}|) + k\ln(1/\lambda)\right)\right)$ rounds, use Hedge to choose $\ell^{(t)}, D^{(t)}$, use the best-response $B_r(\ell^{(t)})$ of Fact 4.1 to set $h^{(t)}$, and sample $z_D^{(t)} \sim D$ from every $D \in \mathcal{D}$ at each iteration to provide a feedback vector to the adversary. Let $p$ be uniform over $h^{(1:T)}$.*

For deterministic multi-calibration, requiring conditional guarantees similarly results in a linear increase in oracle complexity bounds. Here, the oracle is maximizing over both distributions and objectives: $\max_{D\in\mathcal{D},\ell\in\mathcal{G}} \mathcal{L}_{D,\ell}(h)$, where $h$ is the hypothesis it is given.

**Corollary 5.8.** *For any $k$-class multi-calibration problem $(\mathcal{S}, \varepsilon, \lambda)$, running Corollary 4.7's algorithm but using the multi-objective problem $(\mathcal{D}, \mathcal{G}, \mathcal{H})$ of Fact 5.6 results in a deterministic $k$-class predictor $h^* : \mathcal{X} \to [0,1]^k$ such that w.p. $1-\delta$, $h^*$ is $(\mathcal{S}, \varepsilon, \lambda)$-conditionally multi-calibrated predictor. Moreover, only $O\left(\varepsilon^{-2}|\mathcal{S}|\ln(k)\right)$ calls to an agnostic learning oracle are made; this implies, by Lemma 3.5, that only $\widetilde{O}\left(\left(|\mathcal{S}|^{3/2} k\ln(1/\lambda\delta)\right)/\varepsilon^3\right)$ samples are needed overall.*

**Other group fairness notions.** In agnostic multi-objective learning problems, the trade-off between objectives is arbitrated by the worst-off objectives $\max_{\ell^*\in\mathcal{G}} \ell^*(\cdot)$. However, this approach to negotiating trade-offs may be suboptimal when some objectives are inherently more difficult. Specifically, in some multi-objective problems, the non-realizability of the overall problem may be due to the non-realizability of individual objectives rather than trade-offs between objectives. In such cases, it makes sense to allow practitioners to hand-design alternative hypotheses and require our algorithm to try to beat said alternatives, requiring that there is no group for which there exists a competitor $h \in \mathcal{H}^*$ that performs significantly better.

**Definition 5.9.** *For a multi-objective learning problem $(\mathcal{D}, \mathcal{G}, \mathcal{H})$, a solution that is $\varepsilon$-*competitive *w.r.t. a class $\mathcal{H}'$ is a hypothesis $p$ satisfying, for every $i \in \mathcal{I}$*

$$\mathcal{L}^*_{\mathcal{H}'}(p) - \min_{h^*\in\mathcal{H}} \mathcal{L}^*_{\mathcal{H}'}(h^*) \le \varepsilon \text{ where } \mathcal{L}^*_{\mathcal{H}'}(h) \coloneqq \max_{D\in\mathcal{D}} \max_{\ell\in\mathcal{G}} \mathcal{L}_{D,\ell}(h) - \min_{h^*\in\mathcal{H}'} \mathcal{L}_{D,\ell}(h)$$

Only a simple modification is needed to handle competition: amplify your original objectives $\mathcal{G}$ into a new objective set $\mathcal{G}' \coloneqq \bigcap_{h'\in\mathcal{H}'} \{\ell(\cdot) - \ell(h') \mid \ell \in \mathcal{G}\}$ and solve as usual.

**Fact 5.10** (Competitive Multi-Obj Reduction)**.** *Consider a multi-objective learning problem $(\mathcal{D}, \mathcal{G}, \mathcal{H})$. For some choice of $\mathcal{H}'$, let $\mathcal{G}' \coloneqq \bigcap_{h'\in\mathcal{H}'} \{\ell(\cdot) - \ell(h') \mid \ell \in \mathcal{G}\}$. Any solution $p$ that is $\varepsilon$-optimal for the multi-objective learning problem $(\mathcal{D}, \mathcal{G}', \mathcal{H})$, is also $\varepsilon$-competitive w.r.t. $\mathcal{H}'$ for the original problem $(\mathcal{D}, \mathcal{G}, \mathcal{H})$.*

When $\mathcal{H}' = \mathcal{H}$, we can see that competition becomes similar to a *objective-wise* notion of optimality, where a hypothesis $p$ satisfies, for every $\ell \in \mathcal{G}$, that $\mathcal{L}_{D,\ell}(p) - \min_{h^*\in\mathcal{H}} \mathcal{L}_{D,\ell}(h^*) \le \varepsilon$. The following is an immediate consequence of standard multi-distribution learning sample complexity bounds (see Haghtalab et al. [Theorem 4.1 16]).

**Corollary 5.11.** *The following algorithm takes $O\left(\left(\log(|\mathcal{G}|\,|\mathcal{H}|) + |\mathcal{D}|\ln(|\mathcal{D}|/\delta)\right)/\varepsilon^2\right)$ samples and with probability at least $1-\delta$ yields an objective-wise $\varepsilon$-optimal hypothesis for the multi-objective problem $(\{D\}, \mathcal{G}, \mathcal{H})$ when objective-wise optimality is attainable: create a new problem $(\mathcal{D}, \mathcal{G}', \mathcal{H})$ where $\mathcal{G}' \coloneqq \bigcap_{h'\in\mathcal{H}} \{\ell(\cdot) - \ell(h') \mid \ell \in \mathcal{G}\}$, simultaneously run no-regret dynamics by implementing Hedge on $\mathcal{G}'$ and on $\mathcal{H}$ to get $(\ell, h)^{(1:T)}$, and return a uniform distribution over $h^{(1:T)}$.*

Rothblum and Yona [26] study a specific case of this problem where only a single data distribution is involved ($\mathcal{D} = \{D\}$) and where objective-wise optimality is assumed to be attainable. For this "group-compatible" setting, Tosh and Hsu [28] attained optimal rates of $\log(|\mathcal{G}|\,|\mathcal{H}|)/\varepsilon^2$ using a reduction to sleeping experts, which Corollary 5.11 matches with simple no-regret dynamics and extends to multiple distributions.

We further note that not only does Corollary 5.11 attain the minimax sample complexity rate for objective-wise optimal learning, but also requiring objective-wise optimality does not change the minimax rate for the problem of multi-distribution multi-objective learning. This is because there is a sample complexity lower bound of $\widetilde{\Theta}\left((\log(|\mathcal{G}|\,|\mathcal{H}|) + |\mathcal{D}|)/\varepsilon^2\right)$ for basic multi-objective learning (see Haghtalab et al. [16]'s Theorem 4.3). Adding competitiveness or objective-wise optimality requirements to multi-objective learning therefore does not change things significantly: the same algorithms still work and minimax rates do not change beyond constants.

## 6 Acknowledgments

This work was supported in part by the National Science Foundation under grant CCF-2145898, a C3.AI Digital Transformation Institute grant, a Berkeley AI Research (BAIR) Commons grant, and the Mathematical Data Science program of the Office of Naval Research. This work was partially done while Haghtalab and Zhao were visitors at the Simons Institute for the Theory of Computing.

## References

[1] R. Bassily, K. Nissim, A. D. Smith, T. Steinke, U. Stemmer, and J. R. Ullman. Algorithmic stability for adaptive data analysis. *SIAM Journal of Computing*, 50(3), 2021.

[2] A. Blum and T. Lykouris. Advancing subgroup fairness via sleeping experts. In T. Vidick, editor, *11th Innovations in Theoretical Computer Science Conference*, 2020.

[3] A. Blum, N. Haghtalab, A. D. Procaccia, and M. Qiao. Collaborative PAC learning. In I. Guyon, U. von Luxburg, S. Bengio, H. M. Wallach, R. Fergus, S. V. N. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 2392–2401, 2017.

[4] J. Chen, Q. Zhang, and Y. Zhou. Tight bounds for collaborative PAC learning via multiplicative weights. In S. Bengio, H. M. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31*, pages 3602–3611, 2018.

[5] A. P. Dawid. The well-calibrated Bayesian. *Journal of the American Statistical Association*, 77(379):605–610, 1982.

[6] C. Dwork, M. P. Kim, O. Reingold, G. N. Rothblum, and G. Yona. Outcome indistinguishability. In S. Khuller and V. V. Williams, editors, *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1095–1108. ACM, 2021.

[7] C. Dwork, D. Lee, H. Lin, and P. Tankala. New insights into multi-calibration, Jan. 2023.

[8] D. Foster and R. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21:40–55, 1997.

[9] D. P. Foster and S. M. Kakade. Calibration via regression. In G. Seroussi and A. Viola, editors, *Proceedings of 2006 IEEE Information Theory Workshop*, pages 82–86. IEEE, 2006.

[10] D. P. Foster and R. V. Vohra. Asymptotic calibration. *Biometrika*, 85(2):379–390, 1998.

[11] Y. Freund and R. E. Schapire. Game theory, on-line prediction and boosting. In A. Blum and M. J. Kearns, editors, *Proceedings of the Ninth Annual Conference on Computational Learning Theory*, pages 325–332. ACM, 1996.

[12] Y. Freund and R. 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.

[13] P. Gopalan, A. T. Kalai, O. Reingold, V. Sharan, and U. Wieder. Omnipredictors. In M. Braverman, editor, *13th Innovations in Theoretical Computer Science Conference*, pages 79:1–79:21, 2022.

[14] P. Gopalan, M. P. Kim, M. Singhal, and S. Zhao. Low-degree multicalibration. In P.-L. Loh and M. Raginsky, editors, *Conference on Learning Theory*, pages 3193–3234. PMLR, 2022.

[15] V. Gupta, C. Jung, G. Noarov, M. M. Pai, and A. Roth. Online multivalid learning: Means, moments, and prediction intervals. In M. Braverman, editor, *13th Innovations in Theoretical Computer Science Conference*, pages 82:1–82:24, 2022.

[16] N. Haghtalab, M. I. Jordan, and E. Zhao. On-demand sampling: Learning optimally from multiple distributions. In S. Bengio, H. M. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, 2022.

[17] S. Hart. Calibrated forecasts: The minimax proof, 2022.

[18] U. Hebert-Johnson, M. P. Kim, O. Reingold, and G. N. Rothblum. Multicalibration: Calibration for the (computationally-identifiable) masses. In J. G. Dy and A. Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, pages 1944–1953. PMLR, 2018.

[19] C. Jung, C. Lee, M. M. Pai, A. Roth, and R. Vohra. Moment multicalibration for uncertainty estimation. In M. Belkin and S. Kpotufe, editors, *Conference on Learning Theory*, pages 2634–2678. PMLR, 2021.

[20] C. Jung, G. Noarov, R. Ramalingam, and A. Roth. Batch multivalid conformal prediction, 2022.

[21] M. J. Kearns, S. Neel, A. Roth, and Z. S. Wu. Preventing fairness gerrymandering: Auditing and learning for subgroup fairness. In J. G. Dy and A. Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, pages 2569–2577. PMLR, 2018.

[22] M. P. Kim, A. Ghorbani, and J. Y. Zou. Multiaccuracy: Black-box post-processing for fairness in classification. In V. Conitzer, G. K. Hadfield, and S. Vallor, editors, *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pages 247–254. ACM, 2019.

[23] M. Mohri, G. Sivek, and A. T. Suresh. Agnostic federated learning. In K. Chaudhuri and R. Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, pages 4615–4625. PMLR, 2019.

[24] H. L. Nguyen and L. Zakynthinou. Improved algorithms for collaborative PAC learning. In S. Bengio, H. M. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, pages 7642–7650, 2018.

[25] G. Noarov, M. M. Pai, and A. Roth. Online multiobjective minimax optimization and applications, 2021.

[26] G. N. Rothblum and G. Yona. Multi-group agnostic PAC learnability. In M. Meila and T. Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, pages 9107–9115. PMLR, 2021.

[27] S. Sagawa, P. W. Koh, T. B. Hashimoto, and P. Liang. Distributionally robust neural networks. In *6th International Conference on Learning Representations, Conference Track Proceedings*, 2020.

[28] C. J. Tosh and D. Hsu. Simple and near-optimal algorithms for hidden stratification and multi-group learning. In K. Chaudhuri, S. Jegelka, L. Song, C. Szepesvári, G. Niu, and S. Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, pages 21633–21657. PMLR, 2022.

# A Proofs for Section 2

In this section, we restate the technical results mentioned in Section 2 along with references. These are all existing results and we refer interested readers to the references for details.

Lemma 2.1 is a special case of the following lemma about the martingale losses of online learning. We believe this result is folklore, and refer readers interested in an explicit proof to Lemma B.1 of Haghtalab et al. [16], who prove the result in their Fact B.3.

**Lemma A.1** (Haghtalab et al. [16]). *Let $(h, \ell, z)^{(1:T)}$ be the result of running an online learning algorithm on any adversarially chosen sequence $a_{-i}^{(1:T)}$ which we constrain to be a martingale with respect to $\widehat{\phi}^{(1:T)}$. Then, with probability at least $1-\delta$,*

$$Reg_i(a^{(1:T)}) \leq \mathbb{E}\left[\widehat{Reg_i}\left(a^{(1:T)}\right)\right] + O\left(\sqrt{T\ln(|\mathcal{A}_i|/\delta)}\right).$$

Lemma 3.5 originates from adaptive data analysis, and is taken in its current form from Corollary 6.4 in Bassily et al. [1]. This bounds the sample complexity of implementing a noisy-max oracle.

**Lemma 3.5** (Adaptive Data Analysis [1]). *There is an algorithm that for any distribution $D$, any set of loss functions $\mathcal{G} : \mathcal{H} \times \mathcal{X} \to [0,1]$, and any adaptive sequence of hypotheses $h^{(1:T)}$, with probability at least $1-\delta$ returns $\ell^{(1:T)}$ such that for all $t \in [T]$, $\mathcal{L}^{(t)}(h^{(t)}) \geq \max_{\ell^* \in \mathcal{G}} \mathcal{L}^*(h^{(t)}, z) - \varepsilon$, using no more samples from $D$ than:*

$$O\left(\frac{\sqrt{T}}{\varepsilon^2}\log\left(\frac{|\mathcal{G}|}{\varepsilon}\right)\log^{3/2}\left(\frac{1}{\varepsilon\delta}\right)\right) \approx \widetilde{O}\left(\frac{\sqrt{T}\ln(|\mathcal{G}|/\delta)}{\varepsilon^2}\right).$$

# B Proofs of Section 3

In this section, we will prove generalizations of Lemma 3.1 and Lemma 3.2 that directly imply them. As we will be working in a general multi-player setting, we will use the following definitions of regret and weak regret analogous to those of the two-player games setting:

$$\mathrm{Reg}_{\mathbf{A}_a}(\{h_{1:b}^{(t)}, \ell_{1:b}^{(t)}\}_{t\in[T]}) := \max_{\ell_a^* \in \mathcal{G}_a} \sum_{t=1}^{T} \mathop{\mathbb{E}}_{z\sim D}\left[\ell_a^*(h^{(t)}, z)\right] - \sum_{t=1}^{T} \mathop{\mathbb{E}}_{z\sim D}\left[\ell_a^{(t)}(h^{(t)}, z)\right]$$

$$\mathrm{WReg}_{\mathbf{L}_a}^{\mathbf{A}_a}(\{h_{1:b}^{(t)}, \ell_{1:b}^{(t)}\}_{t\in[T]}) := \sum_{t=1}^{T} \mathop{\mathbb{E}}_{z\sim D}\left[\ell_a^{(t)}(h^{(t)}, z)\right] - \min_{h_a^* \in \mathcal{H}_a} \max_{\ell_a^* \in \mathcal{G}_a} \sum_{t=1}^{T} \mathop{\mathbb{E}}_{z\sim D}\left[\ell_a^*(h_a^*, h_{-a}^{(t)}, z)\right]$$

**Lemma B.1** (No-Regret vs. No-Regret (Generalization of Lemma 3.1)). *Let $(h_{1:b}, \ell_{1:b})^{(1:T)}$ be a play history between $b$ learners and $b$ adversaries. Suppose, for every $a \in [b]$, learner $a$ has bounded weak regret $T\varepsilon \geq WReg_{\mathbf{L}_a}^{\mathbf{A}_a}\left((h,\ell)^{(1:T)}\right)$ and $T \cdot \varepsilon \geq Reg_{\mathbf{A}_a}\left((h,\ell)^{(1:T)}\right)$. Then a uniform distribution on $h^{(1:T)}$ is a component-wise $2\varepsilon$-optimal solution for the multi-objective learning problem $(\mathcal{G}, \mathcal{H})$.*

*Proof.* By the regret guarantees of adversaries $\mathbf{A}_a$, we have that for all $a \in [u]$

$$\max_{\ell_a^* \in \mathcal{G}_a} \sum_{t=1}^{T} \mathop{\mathbb{E}}_{z\sim D}\left[\ell_a^*(h^{(t)}, z)\right] - \sum_{t=1}^{T} \mathop{\mathbb{E}}_{z\sim D}\left[\ell_a^{(t)}(h^{(t)}, z)\right] \leq T\varepsilon.$$

Furthermore, by the weak regret guarantees of every learner with respect to the adversaries, we have

$$\sum_{t=1}^{T} \mathop{\mathbb{E}}_{z\sim D}\left[\ell_a^{(t)}(h^{(t)}, z)\right] - \min_{h_a^*\in\mathcal{H}_a} \max_{\ell_a^*\in\mathcal{G}_a} \sum_{t=1}^{T} \mathop{\mathbb{E}}_{z\sim D}\left[\ell_a^*(h_a^*, h_{-a}^{(t)}, z)\right] \leq T\varepsilon.$$

Therefore, for all $a \in [u]$ we have

$$\max_{\ell_a^*} \sum_{t=1}^{T} \mathop{\mathbb{E}}_{z\sim D}\left[\ell_a^*(h_{1:r}^{(t)}, z)\right] - \min_{h_a^*\in\mathcal{H}_a} \max_{\ell_a^*} \sum_{t=1}^{T} \mathop{\mathbb{E}}_{z\sim D}\left[\ell_a^*(h_a^*, h_a^{(t)})\right] \leq 2T\varepsilon.$$

□

**Lemma B.2** (No-regret vs. Best-Response (Generalization of Lemma 3.2))**.** *Let $(h_{1:b}, \ell_{1:b})^{(1:T)}$ be a play history between $b$ learners and $b$ adversaries. Suppose for every $a \in [b]$, learner $a$ has bounded weak regret $T\varepsilon \geq WReg_{\mathbf{L}_a}^{\mathbf{A}_a}\left((h,\ell)^{(1:T)}\right)$, where adversary $a$ has been—on average—best responding: $\sum_{t=1}^{T} \mathcal{L}_a^{(t)}(h^{(t)}) + \varepsilon \geq c \sum_{t=1}^{T} \max_{\ell_a^*} \mathcal{L}_a^*(h^{(t)})$ for some constant $c \in [0,1]$. Then there exists a $t \in [T]$ where $h^{(t)}$ is a component-wise $\frac{1}{c}(2\varepsilon b + (1-c)\alpha)$-optimal solution for the multi-objective learning problem $(\mathcal{D}, \mathcal{G}, \mathcal{H})$, where*

$$\alpha \coloneqq \sum_{a\in[b]} \min_{h_a^*\in\mathcal{H}_a} \max_{\ell_a^*} \sum_{t=1}^{T} \mathcal{L}_a^*(h_a^*, h_{-a}^{(t)}).$$

*Proof.* We want to show that there is a $t \in [T]$ where, for every $a \in [b]$, $\max_{\ell_a^*} \mathcal{L}_a^*(h^{(t)}) \leq \min_{h_a^*\in\mathcal{H}_a} \max_{\ell_a^*} \mathcal{L}_a^*(h_a^*, h_{-a}^{(t)}) + \varepsilon'$. We will use a proof by contradiction. Accordingly, suppose that for all $t \in [T]$, we have that $\exists a \in [b]$ where,

$$\max_{\ell_a^*} \mathcal{L}_a^*(h^{(t)}) > \min_{h_a^*\in\mathcal{H}_a} \max_{\ell_a^*} \mathcal{L}_a^*(h_a^*, h_{-a}^{(t)}) + \varepsilon'.$$

Note that we still always have the general fact that, for any $a \in [r]$,

$$\max_{\ell_a^*} \mathcal{L}_a^*(h^{(t)}) \geq \min_{h_a^*\in\mathcal{H}_a} \max_{\ell_a^*} \mathcal{L}_a^*(h_a^*, h_{-a}^{(t)}).$$

Now take the assumption that the learners are weakly no-regret to see that,

$$\sum_{a\in[b]} \sum_{t=1}^{T} \mathcal{L}_a^{(t)}(h^{(t)}) - \sum_{a\in[b]} \min_{h_a^*\in\mathcal{H}_a} \max_{\ell_a^*} \sum_{t=1}^{T} \mathcal{L}_a^*(h_a^*, h_{-a}^{(t)}) \leq T\varepsilon b.$$

Observing that adversaries are best-responding,

$$\sum_{t=1}^{T} \mathcal{L}_a^{(t)}(h^{(t)}) \geq \sum_{t=1}^{T} c \cdot \max_{\ell_a^*} \mathcal{L}_a^*(h^{(t)}) - \varepsilon,$$

we see that, for any particular $a$:

$$T\varepsilon b \geq \sum_{t=1}^{T} \sum_{a\in[b]} \left(c \cdot \max_{\ell_a^*} \mathcal{L}_a^*(h^{(t)}) - \varepsilon\right) - \sum_{a\in[b]} \min_{h_a^*\in\mathcal{H}_a} \max_{\ell_a^*} \sum_{t=1}^{T} \mathcal{L}_a^*(h_a^*, h_{-a}^{(t)})$$

$$> \sum_{t=1}^{T} c\varepsilon' + \sum_{a\in[b]} \left( c \cdot \min_{h_a^*\in\mathcal{H}_a} \max_{\ell_a^*} \mathcal{L}_a^*(h_a^*, h_{-a}^{(t)}) - \varepsilon \right) - \sum_{a\in[b]} \min_{h_a^*\in\mathcal{H}_a} \max_{\ell_a^*} \sum_{t=1}^{T} \mathcal{L}_a^*(h_a^*, h_{-a}^{(t)})$$

This implies

$$T\left(2\varepsilon b - c\varepsilon'\right) > \sum_{a\in[b]} c \cdot \sum_{t=1}^{T} \min_{h_a^*\in\mathcal{H}_a} \max_{\ell_a^*} \mathcal{L}_a^*(h_a^*, h_{-a}^{(t)}) - \underbrace{\sum_{a\in[b]} \min_{h_a^*\in\mathcal{H}_a} \max_{\ell_a^*} \sum_{t=1}^{T} \mathcal{L}_a^*(h_a^*, h_{-a}^{(t)})}_{\alpha},$$

or in terms of $\alpha$, reaching the contradiction:

$$2\varepsilon b + (1-c)\alpha > c\varepsilon'.$$

□

**Lemma 3.4.** *Take any set of hypotheses $p^{(1:T)}$ where at least one choice of $p^{(t)}$ is $\varepsilon$-optimal. The following procedure* $\textsc{Find}(\mathcal{H}, \mathcal{G})$ *identifies a $2\varepsilon$-optimal solution with probability at least $1-\delta$, and with only $N = O\left(\varepsilon^{-2}\ln(T\left|\mathcal{G}\right|/\delta)\right)$ samples: return $p^{(t^*)}$ where $t^* = \arg\min_{t\in[T]} \max_{\ell\in\mathcal{G}} \widehat{\mathcal{L}}_S(p^{(t)})$ and $S$ is a set of $N$ samples drawn i.i.d. from $D$. This can also be implemented with $T$ calls to an agnostic learning oracle over $\mathcal{G}$ and 1 call to an agnostic learning oracle over $[T]$.*

*Proof.* By the size of $N$, we know that with probability at least $1-\delta$, for all $t\in[T], \ell\in\mathcal{G}$:

$$\left|\widehat{\mathcal{L}}_S(h^{(t)}) - \mathcal{L}_S(h^{(t)})\right| \le \varepsilon/2.$$

For the oracle-efficient implementation, use an agnostic learning oracle to find, for each $t\in[T]$, a $(\ell^*)^{(t)} \in \mathcal{G}$ such that $(\mathcal{L}^*)^{(t)}(h^{(t)}) \ge \max_{\ell^*\in\mathcal{G}} \mathcal{L}^*(h^{(t)}) - \varepsilon$. Then, use another oracle to find $t^*\in[T]$ such that $(\mathcal{L}^*)^{(t^*)}(h^{(t^*)}) \le \min_{t\in[T]} (\mathcal{L}^*)^{(t)}(h^{(t)}) + \varepsilon$. □

# C Proofs for Section 4

For Fact 4.1, we refer interested readers to Hart [17] for an explicit construction of $B_r$ and further discussion. We will sketch an existence proof here to give intuition.

**Fact 4.1** (Hart [17])**.** *Consider any online learning problem where an adversary chooses $\ell(h,(x,y)) = \langle f_\ell(h,x), h(x) - g(y)\rangle$ from a set of linear losses $\mathcal{G}$ and nature chooses datapoints $(x,y)$. Then for any mixed strategy of the adversary $q\in\Delta\mathcal{G}$ and any $r\in\mathbb{N}$, there is an explicitly constructable mapping $B_r : \Delta\mathcal{G} \to (\Delta\mathcal{Y})^{\mathcal{X}}$ that provides a (near-)zero regret strategy for the learner against any datapoint that nature may draw: $\forall z\in\mathcal{Z} : q(B_r(q), z) \le 1/r$. Moreover, $B_r$ is distribution-free and can be computed without samples.*

*Proof Sketch.* To observe the existence of $B_r : \Delta\mathcal{G} \to (\Delta\mathcal{Y})^{\mathcal{X}}$ satisfying $\forall z\in\mathcal{Z} : q(B_r(q), z) \le 1/r$, discretize the output of $B_r$ to intervals of length $1/r$ and define each $B_r(x)$ to be a mixed strategy over their midpoints. Conditioning on an $x\in\mathcal{X}$, consider the game between a learner choosing a strategy for $B_r(x)$ and an adversary choosing a distribution over labels $g(y)$. If the adversary was forced to move first, the learner could choose $B_r(x)$ to be a point distribution on the midpoint closest to $g(y)$, where $|g(y) - B_r(x)| \le 1/r$. By Sion's minimax theorem, there must also be a $1/r$-regret strategy $\widehat{y}$ for the learner if they were to move first; accordingly define $B_r(x) = \widehat{y}$. □

Lemma C.1 is a generalization of Lemma 4.2 to a general online learning setting involving multiple classes, multiple labels (groups of classes), multiple groups, and for a more general class of semi-linear symmetric multi-objective learning problems. For example, this lemma is needed for agnostic multi-calibration and moment multi-calibration. Lemma 4.2 arises as a special case where $r = 1$. Below, let $M \circ M'$ denote Hadamard product.

**Lemma C.1.** *Consider any multi-label online learning problem with loss function $\ell$, a learner which chooses from $\mathcal{H} = (\Delta_k^r)^{\mathcal{X}}$ and an adaptive adversary which chooses $b$-dimensional (say, real-valued) parameters $\beta \in \mathbb{R}^b$. For ease, we will write labels and predictions as matrices of size $r \times k$.*

- Assumption 1 (Linearity) *There is a function $M : \mathcal{H} \times \mathcal{X} \times \{0,1\}^{[u]} \times \mathbb{R}^b \to [-1,1]_{r\times k}$ and $g : \mathcal{Y} \to \Delta_k^r$ s.t. $\ell$ takes the linear form: $\ell(h,(x,y,w),\beta,\gamma) = M(h,x,w,\beta) \circ (h(x) - g(y))$.*
- Assumption 2 (Symmetry) *There exists an $i \in [b]$ s.t. for any $\beta$ and permutation matrix $P$ of size $r \times r$, there exists another $\beta'$ such that $\beta_{-i} = \beta'_{-i}$ and $M(\cdot,\beta) = M(\cdot,\beta')P$.*

*Then, there is an explicit online learning algorithm for the learner such that the following weak regret guarantee holds against any evaluation distribution $D$ over $\mathcal{X} \times \mathcal{Y} \times \{0,1\}^{[u]}$:*

$$WReg^i \coloneqq \sum_{t=1}^{T} \mathcal{L}_{w^{(t)}}(h^{(t)},\beta^{(t)}) - \min_{h^*\in\mathcal{H}} \max_{\beta_i^*\in\mathbb{R}} \sum_{t=1}^{T} \mathcal{L}_{w^{(t)}}(h^*, \beta_{-i}^{(t)}\beta_i^*) \le O\left(\sqrt{r\ln(k)T} + \Delta\right),$$

*where $\Delta$ is a covariance adjustment $\Delta \coloneqq -\mathbb{E}_x\left[Cov\left[\sum_{t=1}^{T} M(h^{(t)},x,w,\beta^{(t)}),\ g(y) \mid x\right]\right]$ if $y$ is correlated with $w$ and 0 otherwise.*

*Proof.* We will construct our online learning algorithm by defining each iterate $h^{(t)}$ in terms of its behavior on each $x$ in the domain $\mathcal{X}$. For each $x$, we interpret each prediction $h^{(t)}(x)$ as a $r$ distributions over $k$ imaginary actions returned by running a no-regret algorithm on an adaptive sequence of losses $\gamma^{(1)},\ldots,\gamma^{(t-1)}$ where $\gamma^{(\tau)} : [k]^r \to [-1,1]$. In particular, we consider the loss sequence $\gamma^{(\tau)}(\widehat{y}) \coloneqq \mathbb{E}_{w\sim D}\left[M(h^{(\tau)},x,w,\beta^{(\tau)}) \circ \widehat{y} \mid x\right]$ for all $\widehat{y} \in [k]^r$. We note that $\gamma_x^{(\tau)}$ is a valid loss function for adaptive adversaries, who can observe the players mixed strategy $h^{(t)}(x)$ for all $x$, but not the action $\widehat{y}$ realized from the strategies. Using Hedge, or any other online algorithm with similar regret for $k^r$ choices of $\widehat{y}$, we have that,

$$\sum_{t=1}^{T} \mathbb{E}_{\widehat{y}\sim h^{(t)}}\left[\gamma_x^{(t)}(\widehat{y})\right] \le \min_{y^*\in[k]^r} \sum_{t=1}^{T} \gamma_x^{(t)}(y^*) + O\left(\sqrt{r\ln(k)T}\right).$$

Replacing in the definition of $\gamma_x^{(t)}$ and leveraging its linearity to substitute $\widehat{y}^{(t)} \sim h^{(t)}$ with the probability vector $h^{(t)}$ we have,

$$\sum_{t=1}^{T} \mathbb{E}_{w\sim D}\left[M(h^{(t)},x,w,\beta^{(t)}) \mid x\right] \circ h^{(t)} \le \min_{y^*\in[k]^r} \sum_{t=1}^{T} \mathbb{E}_{w\sim D}\left[M(h^{(t)},x,w,\beta^{(t)}) \circ y^* \mid x\right] + O\left(\sqrt{r\ln(k)T}\right).$$

Subtracting from both sides $C \coloneqq \mathbb{E}_{w,y\sim D}\left[\sum_{t=1}^{T} M(h^{(t)},x,w,\beta^{(t)}) \circ g(y) \mid x\right]$, for any $\eta \in \mathcal{Y}^{\mathcal{X}}$ we have:

$$\mathbb{E}_{x,y,w\sim D}\left[\sum_{t=1}^{T} M(h^{(t)},x,w^{(t)},\beta^{(t)}) \circ (h^{(t)} - g(y^{(t)}))\right]$$

$$\leq \mathop{\mathbb{E}}_{x\sim D}\Bigg[\underbrace{\min_{y^*\in[k]^r}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{w\sim D}\Big[M(h^{(t)},x,w,\beta^{(t)})\mid x\Big]\circ(y^*-\eta(x))\mid x\Big]}_{(a)}\Bigg]+O\left(\sqrt{r\ln(k)T}\right)$$

$$+\underbrace{\mathop{\mathbb{E}}_{x\sim D}\Bigg[\sum_{t=1}^{T}\mathop{\mathbb{E}}_{(w,y)\sim D}\Big[M(h^{(t)},x,w,\beta^{(t)})\circ(\eta(x)-g(y))\mid x\Big]\Bigg]}_{(b)}.$$

We note that $(a)\leq 0$, since the minimizer can always choose $y^*$ such that $M_{1:T}\circ y^*=C$, where $C=\sum_{t=1}^{T}\mathbb{E}_w\left[M(h^{(t)},x,w,\beta^{(t)})\mid x\right]\circ\eta(x)$. We also note that $(b)=-\mathbb{E}_x\left[\mathrm{Cov}\left[\sum_{t=1}^{T}M(h^{(t)},x,w,\beta^{(t)}),\ g(y)\mid x\right]\right]$ if we choose $\eta(x)$ to be the label distribution of $x$ marginalized over $w$.

What remains to be shown is that $\sum_{t=1}^{T}\max_{\beta_i^*}\mathcal{L}(h^*,\beta_{-i}^{(t)}\beta_i^*)$ is non-negative. We will now leverage our symmetry assumption. Let $\mathcal{P}$ denote the space of all $k\times k$ permutation matrices. Further define $\beta_i^P\in\mathbb{R}$ to be the value implied by Assumption 2 such that $M(\cdot,\beta_{-i}^{(t)}\beta_i^P)=M(\cdot,\beta_{-i}^{(t)}\beta_i^*)P$.

$$\begin{aligned}
&\mathop{\mathbb{E}}_{x}\left[\sum_{P\in\mathcal{P}}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{w,y}\left[M(h^*,x,w,\beta_{-i}^{(t)},\beta_i^P)\circ(h^*(x)-g(y))\mid x\right]\right]\\
&=\sum_{t=1}^{T}\mathop{\mathbb{E}}_{x,w,y}\left[\sum_{P\in\mathcal{P}}M(h^*,x,w,\beta_{-i}^{(t)},\beta_i^*)P\circ(h^*(x)-g(y))\right]\\
&=\sum_{t=1}^{T}\mathop{\mathbb{E}}_{x,w,y}\left[\left(\underbrace{\sum_{P\in\mathcal{P}}M(h^*,x,w,\beta_{-i}^{(t)},\beta_i^*)P}_{\text{Matrix with identical rows}}\right)\circ(h^*(x)-g(y))\right]\\
&=0,
\end{aligned}$$

where the final equality is because $h^*(x)$ and $g(y)$ are right stochastic matrices. Thus, even if

$$\sum_{t=1}^{T}\mathbb{E}\left[M(h^*,x,w,\beta_{-i}^{(t)},\beta_i^I)\circ(h^*(x)-g(y))\right]<0,$$

where $I$ is the identity permutation, there must exists another summand

$$\sum_{t=1}^{T}\mathbb{E}\left[M(h^*,x,w,\beta_{-i}^{(t)},\beta_i^P)\circ(h^*(x)-g(y))\right]>0.$$

□