JMLR: Workshop and Conference Proceedings 19 (2011) 27–46 24th Annual Conference on Learning Theory

Blackwell Approachability and No-Regret Learning are Equivalent

Jacob Abernethy JAKE@CS.BERKELEY.EDU
University of California, Berkeley
Division of Computer Science

Peter L. Bartlett BARTLETT@CS.BERKELEY.EDU
Univ. of California, Berkeley
Division of Computer Science
Department of Statistics

Elad Hazan EHAZAN@IE.TECHNION.AC.IL
Technion - Israel Institute of Technology
Industrial Engineering and Management

Editor: Sham Kakade, Ulrike von Luxburg

Abstract

We consider the celebrated Blackwell Approachability Theorem for two-player games with vector payoffs. Blackwell himself previously showed that the theorem implies the existence of a "no-regret" algorithm for a simple online learning problem. We show that this relationship is in fact much stronger, that Blackwell's result is equivalent to, in a very strong sense, the problem of regret minimization for Online Linear Optimization. We show that any algorithm for one such problem can be efficiently converted into an algorithm for the other. We provide one novel application of this reduction: the first *efficient* algorithm for calibrated forecasting.

1. Introduction

Von Neumann's minimax theorem (1928) establishes a central result in the theory of two-player zero-sum games, essentially by providing a prescription to both players. This prescription is in the form of a pair of optimal strategies, either of which attains the optimal worst-case value of the game even without knowledge of the opponent's strategy. However, the theorem fundamentally requires that both players have utility that can be expressed as a *scalar*. In 1956, in response to von Neumann's result, David Blackwell posed an intriguing question: what guarantee can we hope to achieve when playing a two-player game with a *vector-valued payoff*?

When our payoffs are non-scalar quantities, it does not make sense to ask "can we earn at least x?". A sensible generalization is, "can we guarantee that our vector payoff lies in some convex set S?" In this case the story is more difficult, and Blackwell observed that an oblivious strategy does not suffice—in short, we do not achieve "minimax duality" for vector-payoff games as we can when the payoff is a scalar. Blackwell was able to prove that this negative result applies only for *one-shot games*. In his celebrated Approachability Theorem (Blackwell, 1956), one can achieve a duality statment *in the limit* when the game is played repeatedly, and the player may learn from his opponent's prior actions. Blackwell constructed an algorithm (that is, an adaptive strategy) that guarantees the average payoff vector "approaches" S.

Blackwell's Approachability Theorem has the flavor of learning in repeated games, a topic which has received much interest. In particular, there are a wealth of recent results on so-called *no-regret learning algorithms* for making repeated decisions given an arbitrary (and potentially adversarial) sequence of cost functions. The first no-regret algorithm for a "discrete action" setting was given in a seminal paper by James Hannan in 1956 (Hannan, 1957). That same year, David Blackwell pointed out (Blackwell, 1954) that his Approachability result leads, as a special case, to an algorithm with essentially the same low-regret guarantee proven by Hannan.

Over the years several other problems have been reduced to Blackwell approachability, including asymptotic calibration (Foster and Vohra, 1998), online learning with global cost functions (Even-Dar et al., 2009) and more (Mannor and Shimkin, 2008). Indeed, it has been presumed that approachability, while establishing the existence of a no-regret algorithm, is strictly more powerful than regret-minimization; hence its utility in such a wide range of problems. In the present paper we prove, to the contrary, that Blackwell's Approachability Theorem is equivalent, in a very strong sense, to no-regret learning for the setting of *Online Linear Optimization*. This shows that the connection discovered by Blackwell, between regret and approachability, is much stronger than originally supposed.

More specifically, we show how any no-regret algorithm can be converted into an algorithm for Approachability and vice versa. This algorithmic equivalence is achieved via the use of *conic duality*: an approachability problem over a convex cone K can be reduced to an online linear optimization instance where we must "learn" within the *polar cone* K^0. The reverse direction is similar. This equivalence provides a range of benefits and one such is "asymptotic calibrated forecasting". The calibration problem was reduced to Blackwell's Approachability Theorem by Foster (1999), and a handful of other calibration techniques have been proposed, yet none have provided any efficiency guarantees on the strategy. Using a similar reduction from calibration to approachability, and by carefully constructing the reduction from approachability to online linear optimization, we achieve the first efficient calibration algorithm.

Related work There is by now vast literature on all three main topics of this paper: approachability, online learning and calibration, see (Cesa-Bianchi and Lugosi, 2006) for an excellent exposition.

Calibration is a fundamental notion in prediction theory and has found numerous applications in economics and learning. Dawid (1982) was the first to define calibration, with numerous algorithms later given by Foster and Vohra (1998), Fudenberg and Levine (1999), Hart and Mas-Colell (2000) and more (see e.g. (Sandroni et al., 2003; Perchet, 2009)). Foster has given a calibration algorithm based on approachability (Foster, 1999). There are numerous definitions (mostly asymptotic) of calibration in the literature. In this paper we give precise finite-time rates of calibration. Furthermore, we give the first *efficient* algorithm for calibration: attaining ε-calibration (formally defined later) required a running time of $poly(\frac{1}{\varepsilon})$ for all previous algorithms, whereas our algorithm runs in time proportional to $\log \frac{1}{\varepsilon}$.

2. Game Theory Preliminaries

2.1. Two-Player Games

Formally, a two-player normal-form game is defined by a pair of action sets $[n]$ and $[m]$, for natural numbers n, m, and a pair of utility functions $u_1, u_2 : [n] \times [m] \to \mathbb{R}$. When player 1 chooses action i and player 2 chooses action j, player 1 and player 2 receive utilities $u_1(i, j)$ and $u_2(i, j)$ respectively. An important class of two-player games are known as *zero-sum*, in that $u_1 \equiv -u_2$. For zero-sum games we drop the subcripts on u_1, u_2 and simply write $u(i, j)$ for player 1's utility, and $-u(i, j)$ for player 2's utility. For the remainder of this section, we shall be concerned entirely with zero-sum games, hence we will refer to player 1 as the Player and player 2 as the Adversary.

It is natural to assume that the players in a game may include randomness in their choice of action; simple games such as Rock-Paper-Scissors require randomness to achieve optimality. When the players choose their actions randomly according to the distributions $p \in \Delta_n$ and $q \in \Delta_m$, respectively, the *expected utility* for the Player is $\sum_{i,j} p(i)q(j)u(i, j)$. Von Neumann's minimax theorem, widely considered the first key result in game theory, tells us that both the Player and the Adversary have an "optimal" randomized strategy that can be played without knowledge of the strategy of their respective opponent.

Theorem 1 (Von Neumann's Minimax Theorem (Neumann et al., 1947)) *For any integers $n, m > 0$ and any utility function $u : [n] \times [m] \to \mathbb{R}$,*

$$\max_{p \in \Delta_n} \min_{q \in \Delta_m} \sum_{i,j} p(i)q(j)u(i, j) = \min_{q \in \Delta_m} \max_{p \in \Delta_n} \sum_{i,j} p(i)q(j)u(i, j)$$

The statement of the minimax theorem is often referred to as *duality* as it swaps the min and max. This result can be used to establish strong duality for linear programming. It was proven by Maurice Sion in the 1950's that von Neumann's notion of duality can be extended further, for a much larger class of input spaces and a more general class of functions.

Theorem 2 (Sion (1958)[1]) *Given convex compact sets $\mathcal{X} \subset \mathbb{R}^n, \mathcal{Y} \subset \mathbb{R}^m$, and a function $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ convex and concave in its first and second arguments respectively, we have*

$$\inf_{\mathbf{x} \in \mathcal{X}} \sup_{\mathbf{y} \in \mathcal{Y}} f(\mathbf{x}, \mathbf{y}) = \sup_{\mathbf{y} \in \mathcal{Y}} \inf_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}, \mathbf{y}).$$

In the present work we shall not need anything quite so general, although we use this theorem to generalize slightly the class of two-player zero-sum games. Rather than define the actions of our players as being drawn randomly from discrete sets $[n]$ and $[m]$, let the players' decision space be characterized by given compact convex sets $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{Y} \subset \mathbb{R}^m$ respectively. In addition, we shall assume that the utility is characterized by a *biaffine* function $u : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$; that is, $u(\alpha\mathbf{x} + (1 - \alpha)\mathbf{x}', \mathbf{y}) = \alpha u(\mathbf{x}, \mathbf{y}) + (1 - \alpha)u(\mathbf{x}', \mathbf{y})$ and $u(\mathbf{x}, \alpha\mathbf{y} + (1 - \alpha)\mathbf{y}') = \alpha u(\mathbf{x}, \mathbf{y}) + (1 - \alpha)u(\mathbf{x}, \mathbf{y}')$ for every $0 \le \alpha \le 1$, $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ and $\mathbf{y}, \mathbf{y}' \in \mathcal{Y}$. Following Sion's theorem, we arrive at the following.

Corollary 3 *For compact convex sets $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{Y} \subset \mathbb{R}^m$ and any biaffine function $u : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, we have*

$$\max_{\mathbf{x} \in \mathcal{X}} \min_{\mathbf{y} \in \mathcal{Y}} u(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{y} \in \mathcal{Y}} \max_{\mathbf{x} \in \mathcal{X}} u(\mathbf{x}, \mathbf{y})$$

This alternative description of a zero-sum game has two advantages. First, we now assume that both players are deterministic. That is, we have converted the notion of a randomized strategy on a discrete action space to a deterministic strategy $\mathbf{x}$ inside of a convex set $\mathcal{X}$. Rather than evaluate the expected utility of a randomized action, this expectation is now incorporated via the linearity of $u(\cdot,\cdot)$. Note, crucially, that the assumptions that u is biaffine and $\mathcal{X}$ and $\mathcal{Y}$ are convex imply that neither player gains from randomness, as $\mathbb{E}_{\mathbf{x}}\mathbb{E}_{\mathbf{y}}\, u(\mathbf{x},\mathbf{y}) = u(\mathbb{E}_{\mathbf{x}}\mathbf{x}, \mathbb{E}_{\mathbf{y}}\mathbf{y})$.

A second advantage of this framework is that it allows us to work with action spaces that might seem prohibitively large. For example, we can imagine a game in which each player must select a route in a graph G between two endpoints, and the utility is the amount of overlap of their paths. The set of paths in a graph is exponential, and even counting the number of such paths is $\#P$-hard. However, we may instead set $\mathcal{X}$ and $\mathcal{Y}$ to be the *flow polytope* of G. The flow polytope can be described by a polynomially-sized number of constraints, and hence is much easier to work with.

2.2. Vector-Valued Games

Let us now turn our attention to Blackwell's question: what can be guaranteed when the utility function of the zero-sum game is *vector-valued*? Following the definition in the previous section, we can define a vector-valued game in terms of some biaffine utility function $\mathbf{u} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^d$ from a product of two convex compact decision spaces $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{Y} \subset \mathbb{R}^m$ to d-dimensional space. The biaffine property is defined in the natural way.

Note that we may not apply our usual notions of utility maximization when dealing with vector-valued games—what does it mean to "maximize" a vector? Furthermore, the concept of "zero-sum" is not immediately clear. Blackwell proposed the following framework: suppose that the Player, who selects $\mathbf{x} \in \mathcal{X}$, would like his vector payoff $\mathbf{u}(\mathbf{x},\mathbf{y})$ to land inside of a particular closed convex set $S \subset \mathbb{R}^d$, where S is fixed and known to both players. We shall say that the Player wants to *satisfy* S. The Adversary, who selects $\mathbf{y} \in \mathcal{Y}$, would like to prevent the Player from satisfying S.

Let us return our attention to the simple case of scalar-valued games discussed in Section 2.1. The duality statement achieved in the Minimax Theorem, typically stated in terms of swapping the order of min and max, can instead be formulated in terms of swapping quantifiers $\forall$ and $\exists$.

Proposition 1 *For any convex compact sets $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{Y} \subset \mathbb{R}^m$, and any biaffine utility function $u : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, we have the following implication for any $c \in \mathbb{R}$:*

$$\forall \mathbf{y} \in \mathcal{Y}\ \exists \mathbf{x} \in \mathcal{X} :\ \ u(\mathbf{x},\mathbf{y}) \in [c,\infty) \quad \Longrightarrow \quad \exists \mathbf{x} \in \mathcal{X}\ \forall \mathbf{y} \in \mathcal{Y} :\ \ u(\mathbf{x},\mathbf{y}) \in [c,\infty).$$

This proposition is simply another way to state duality, in the following form:

$$\min_{\mathbf{y}\in\mathcal{Y}} \max_{\mathbf{x}\in\mathcal{X}} u(\mathbf{x},\mathbf{y}) \geq c \quad \Longrightarrow \quad \max_{\mathbf{x}\in\mathcal{X}} \min_{\mathbf{y}\in\mathcal{Y}} u(\mathbf{x},\mathbf{y}) \geq c.$$

Put another way, if the Player can earn c by choosing his strategy *with knowledge of* the Adversary's strategy, then he can earn c obliviously as well.

Here we have simply taken the Minimax Theorem and stated it in terms of satisfying a set, namely the set $S = [c,\infty)$ for some value c. This interpretation begs the question: can we achieve a similar "duality" statement for vector-valued games? In other words, given a biaffine utility function $\mathbf{u} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^d$ and any convex set $S \subset \mathbb{R}^d$, does the statement

$$\forall \mathbf{y} \in \mathcal{Y}\, \exists \mathbf{x} \in \mathcal{X} : \ \mathbf{u}(\mathbf{x}, \mathbf{y}) \in S \quad \implies \quad \exists \mathbf{x} \in \mathcal{X}\, \forall \mathbf{y} \in \mathcal{Y} : \ \mathbf{u}(\mathbf{x}, \mathbf{y}) \in S$$

hold in general? The answer, unfortunately, is *no!* Consider the following easy example: $\mathcal{X} = \mathcal{Y} := [0, 1]$, the payoff is simply $\mathbf{u}(x, y) := (x, y)$ for $x, y \in [0, 1]$, and the set in question is $S := \{(z, z)\ \forall z \in [0, 1]\}$. Certainly the premise is true, since for every y there exists an x, namely $x = y$, such that $\mathbf{u}(x, y) \in S$. On the other hand, there is no such single x for which $\mathbf{u}(x, y) \in S$ for any y.

2.3. Blackwell Approachability

While we might hope that minimax duality, framed in terms of set satisfiability, would extend from scalar-valued games to vector-valued games, the previous example appears to be a nail in the coffin. But in fact the story is not quite so bad: the proposed example is difficult because it is a *one-shot* game. What Blackwell observed, and led to the Approchability Theorem, is that if the game is played *repeatedly* then one can achieve duality "in the limit." To make this precise we introduce some definitions.

Definition 4 *A* Blackwell instance *is a tuple* $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot, \cdot), S)$, *with* $\mathcal{X} \subset \mathbb{R}^n$ *and* $\mathcal{Y} \subset \mathbb{R}^m$ *compact and convex,* $\mathbf{u} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^d$ *biaffine, and* $S \subset \mathbb{R}^d$ *convex and closed. For any instance* $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot, \cdot), S)$, *we say that*

- S *is* satisfiable *if* $\exists \mathbf{x} \in \mathcal{X}\, \forall \mathbf{y} \in \mathcal{Y} : \ \mathbf{u}(\mathbf{x}, \mathbf{y}) \in S$.
- S *is* response-satisfiable *if* $\forall \mathbf{y} \in \mathcal{Y}\, \exists \mathbf{x} \in \mathcal{X} : \ \mathbf{u}(\mathbf{x}, \mathbf{y}) \in S$.
- S *is* halfspace-satisfiable *if, for any halfspace* $H \supseteq S$, H *is satisfiable.*

To recap, when our utility function $\mathbf{u}$ is scalar-valued, i.e. for zero-sum games where $d = 1$, then minimax duality holds and, according to Proposition 1, this be rephrased as "If $S := [c, \infty)$ is response-satisfiable then S is satisfiable." On the other hand, for vector-valued games it is not the case in general that "S is response-satisfiable $\implies$ S is satisfiable" for arbitrary sets S. What Blackwell showed is that response-satisfiability does lead to a weaker condition, termed *approachability.* Before we define this precisely, let us use the notation $\texttt{dist}(\mathbf{z}, U)$ to denote the distance between a point $\mathbf{z}$ and some convex set U, that is $\inf_{\mathbf{x} \in U} \|\mathbf{z} - \mathbf{x}\|$.

Definition 5 *Given a Blackwell instance* $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot, \cdot), S)$, *we say that* S *is* approachable *if there exists some algorithm* $\mathcal{A}$ *which selects points in* $\mathcal{X}$ *such that, for any sequence* $\mathbf{y}_1, \mathbf{y}_2, \ldots \in \mathcal{Y}$, *we have*

$$\texttt{dist}\Big(\tfrac{1}{T} \textstyle\sum_{t=1}^{T} \mathbf{u}(\mathbf{x}_t, \mathbf{y}_t), S\Big) \to 0 \quad \textit{as} \quad T \to \infty,$$

where $\mathbf{x}_t \leftarrow \mathcal{A}(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{t-1})$.

Under this new notion, we now allow the Player to implement an *adaptive strategy* for a repeated version of the game, and we require that the average utility vector becomes arbitrarily close to S. Intuitively, we may think of approachability as "satisfiability in the limit".

Theorem 6 (Blackwell's Approachability Theorem (Blackwell, 1956)) *For any Blackwell instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S)$, S is approachable if and only if it is response-satisfiable.*

The beauty of this theorem is that, while we may not be able to satisfy S in a one-shot version of the game, we can satisfy the set "on average" if we may play the game indefinitely.

This version of the theorem, which appears in Even-Dar et al. (2009), is not the one usually attributed to Blackwell. The original theorem uses the concept of halfspace satisfiability. It is not difficult to establish the equivalence of the two statements via the following lemma, whose proof uses a nice application of minimax duality.

Lemma 7 *For any Blackwell instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S)$, S is response-satisfiable if and only if it is halfspace-satisfiable.*

Proof ($\Longrightarrow$) Assume that S is response-satisfiable. Hence, for any $\mathbf{y}$ there is an $\mathbf{x}_\mathbf{y}$ such that $\mathbf{u}(\mathbf{x}_\mathbf{y}, \mathbf{y}) \in S$. Now take any halfspace $H \supset S$ parameterized by $\boldsymbol{\theta}, c$, that is $H = \{\mathbf{z} : \langle \boldsymbol{\theta}, \mathbf{z} \rangle \le c\}$. Then let us define a scalar-valued game with utility $u(\mathbf{x}, \mathbf{y}) = \langle \boldsymbol{\theta}, \mathbf{u}(\mathbf{x}, \mathbf{y}) \rangle$. Notice that $H \supset S$ implies that $\langle \boldsymbol{\theta}, \mathbf{z} \rangle \le c$ for all $\mathbf{z} \in S$. Since S is response-satisfiable, for every $\mathbf{y}$ there is an $\mathbf{x}_\mathbf{y}$ such that $\mathbf{u}(\mathbf{x}_\mathbf{y}, \mathbf{y}) \in S \implies u(\mathbf{x}_\mathbf{y}, \mathbf{y}) \le c$. We then immediately see that

$$\max_{\mathbf{y} \in \mathcal{Y}} \min_{\mathbf{x} \in \mathcal{X}} u(\mathbf{x}, \mathbf{y}) \le \max_{\mathbf{y} \in \mathcal{Y}} u(\mathbf{x}_\mathbf{y}, \mathbf{y}) \le c.$$

It follows from Corollary 3 that $\min_{\mathbf{x} \in \mathcal{X}} \max_{\mathbf{y} \in \mathcal{Y}} u(\mathbf{x}, \mathbf{y}) \le c$. Let $\mathbf{x}^* \in \mathcal{X}$ be any minimizer of the latter expression and notice that, for any $\mathbf{y} \in \mathcal{Y}$, we have that $u(\mathbf{x}^*, \mathbf{y}) \le c$. It follows immediately that H is satisfiable.

($\Longleftarrow$) Assume that S is not response-satisfiable. Hence, there must exists some $\mathbf{y}_0 \in \mathcal{Y}$ such that $\mathbf{u}(\mathbf{x}, \mathbf{y}_0) \notin S$ for every $\mathbf{x} \in \mathcal{X}$. Consider the set $U := \{\mathbf{u}(\mathbf{x}, \mathbf{y}_0) \text{ for all } \mathbf{x} \in \mathcal{X}\}$ and notice that U is convex since $\mathcal{X}$ is convex and $\mathbf{u}(\cdot, \mathbf{y}_0)$ is affine. Furthermore, because S is convex and $S \cap U = \emptyset$ by assumption, there must exist some halfspace H separating the two sets, that is $S \subseteq H$ and $H \cap U = \emptyset$. By construction, we see that for any $\mathbf{x}$, $\mathbf{u}(\mathbf{x}, \mathbf{y}_0) \notin H$ and hence H is not satisfiable. It follows immediately that S is not halfspace-satisfiable. ■

Although it is not posed in this language, Blackwell's original theorem uses the concept of a *halfspace oracle*. Given a Blackwell instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S)$, define a halfspace oracle to be a function $\mathcal{O}$ that takes as input any halfspace $H \supset S$ and returns a point $\mathcal{O}(H) = \mathbf{x}_H \in \mathcal{X}$, and we shall refer to a halfspace oracle as *valid* if it satisfies that for each halfspace $H \supset S$, $\mathbf{u}(\mathbf{x}_H, \mathbf{y}) \in H$ for any $\mathbf{y} \in \mathcal{Y}$.

Theorem 8 *For any Blackwell instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S)$, the set S is approachable if and only if there exists a valid halfspace oracle.*

Notice that the existence of a valid halfspace oracle is equivalent to the halfspace-satisfiability condition. Hence, via Lemma 7, this theorem is equivalent to Theorem 6.

To achieve approachability, following Definition 5 one must construct an algorithm $\mathcal{A}$ that maps the observed subsequence $\mathbf{y}_1, \ldots, \mathbf{y}_{t-1} \in \mathcal{Y}$ to a point $\mathbf{x}_t \in \mathcal{X}$. By the previous theorem, in order for the set S to be approachable, there must be a valid halfspace oracle $\mathcal{O}$, and hence $\mathcal{A}$ may make calls to $\mathcal{O}$. Blackwell actually provides such an algorithm, quite elegant for its simplicity, which can be found in his original work (Blackwell, 1956) as well as in the book of Cesa-Bianchi and Lugosi (2006).

We note that, when an approachability algorithm $\mathcal{A}$ is adapted to a Blackwell instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S)$, and makes calls to a halfspace oracle $\mathcal{O}$, we may write $\mathcal{A}^{\mathcal{O}}_{\mathcal{X},\mathcal{Y},\mathbf{u},S}$ to make the dependence clear.

3. Online Linear Optimization

Online Convex Optimization (OCO) has become a popular topic within Machine Learning since it was introduced by Zinkevich (2003), and there has been much followup work (Shalev-Shwartz and Singer, 2007; Rakhlin et al., 2010; Hazan, 2010; Abernethy et al., 2009). It provides a generic problem template and was shown to generalize several existing problems in the realm of online learning and repeated decision making. Among these are online pattern classification, the "experts" or "hedge" setting, and sequential portfolio optimization (Freund and Schapire, 1995; Hazan et al., 2007).

In the OCO setting, we imagine an online game between Player and Nature. Assume the Player is given a convex decision set $\mathcal{K} \subset \mathbb{R}^d$ and must make a sequence of a decisions $\mathbf{x}_1, \mathbf{x}_2, \ldots \in \mathcal{K}$. After committing to $\mathbf{x}_t$, Nature reveals a convex loss function ℓ_t, and Player pays $\ell_t(\mathbf{x}_t)$. The performance of the Player is typically measured by *regret* which we shall define below. In the present work we shall be concerned with the more specific problem of Online Linear Optimization (OLO) where the loss functions are assumed to be linear, $\ell_t(\mathbf{x}) = \langle \mathbf{f}_t, \mathbf{x} \rangle$ for some $\mathbf{f}_t \in \mathbb{R}^d$.

We define the Player's adaptive strategy $\mathcal{L}$, which we refer to as an *OLO algorithm*, as a function which takes as input a subsequence of loss vectors $\mathbf{f}_1, \ldots, \mathbf{f}_{t-1}$ and returns a point $\mathbf{x}_t \leftarrow \mathcal{L}(\mathbf{f}_1, \ldots, \mathbf{f}_{t-1})$, where $\mathbf{x}_t \in \mathcal{K}$.

Definition 9 *Given an OLO algorithm $\mathcal{L}$ and a sequence of loss vectors $\mathbf{f}_1, \mathbf{f}_2, \ldots \in \mathbb{R}^d$, let* $\mathrm{Regret}(\mathcal{L}; \mathbf{f}_{1:T}) := \sum_{t=1}^T \langle \mathbf{f}_t, \mathbf{x}_t \rangle - \min_{\mathbf{x} \in K} \sum_{t=1}^T \langle \mathbf{f}_t, \mathbf{x} \rangle$. *When the sequence of loss vectors is clear, we may simply write* $\mathrm{Regret}_T(\mathcal{L})$.

An important question is whether an OLO algorithm has a regret rate which scales *sublinearly* in T. A sublinear regret is key, for then our average performance, in the long run, is essentially no worse than the best in hindsight. We use the term *no-regret* algorithm when it possesses this property.

Theorem 10 *For any bounded decision set $\mathcal{K} \subset \mathbb{R}^d$ there exists an algorithm $\mathcal{L}_{\mathcal{K}}$ such that* $\mathrm{Regret}_T(\mathcal{L}_{\mathcal{K}}) = o(T)$ *for any sequence of loss vectors $\{\mathbf{f}_t\}$ with bounded norm.*

Later in the paper we provide one such algorithm, known as Online Gradient Descent, proposed by Zinkevich (2003).

Before proceeding, let us demonstrate the value of no-regret algorithms by proving an aforementioned result. We shall sketch a proof of the minimax statement of Corollary 3. Assume we are given convex and compact decision space $\mathcal{X} \subset \mathbb{R}^n$ and $\mathcal{Y} \subset \mathbb{R}^m$, and without loss of generality assume we have a utility function $u : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ of the form $u(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top M \mathbf{y}$ for some $M \in \mathbb{R}^{n \times m}$. Weak duality, i.e. $\min_{\mathbf{y} \in \mathcal{Y}} \max_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\top M \mathbf{y} \geq \max_{\mathbf{x} \in \mathcal{X}} \min_{\mathbf{y} \in \mathcal{Y}} \mathbf{x}^\top M \mathbf{y}$ is trivial, and so we turn our attention to the reverse inequality. We shall imagine our game is played repeatedly, where on round t the first player chooses

$\mathbf{x}_t$ and the second chooses $\mathbf{y}_t$, but where both players select their strategies according to a no-regret algorithm. For every t we shall set $\mathbf{x}_t \leftarrow \mathcal{L}_{\mathcal{X}}(\mathbf{f}_1, \ldots, \mathbf{f}_{t-1})$ and $\mathbf{y}_t \leftarrow \mathcal{L}_{\mathcal{Y}}(\mathbf{g}_1, \ldots, \mathbf{g}_{t-1})$, where we define the vectors $\mathbf{f}_t := -M\mathbf{y}_t$ and $\mathbf{g}_t^\top := \mathbf{x}_t^\top M$. By applying the definition of regret twice, we have

$$\frac{1}{T}\sum_{t=1}^T \mathbf{x}_t^\top M\mathbf{y}_t = \min_{\mathbf{y}\in\mathcal{Y}} \left(\frac{1}{T}\sum_{t=1}^T \mathbf{x}_t\right)^\top M\mathbf{y} + \frac{\text{Regret}_T(\mathcal{L}_{\mathcal{Y}})}{T} \quad\leq\quad \max_{\mathbf{x}\in\mathcal{X}}\min_{\mathbf{y}\in\mathcal{Y}} \mathbf{x}^\top M\mathbf{y} + \frac{o(T)}{T}, \quad (1)$$

$$\frac{1}{T}\sum_{t=1}^T \mathbf{x}_t^\top M\mathbf{y}_t = \max_{\mathbf{x}\in\mathcal{X}} \mathbf{x}^\top M\left(\frac{1}{T}\sum_{t=1}^T \mathbf{y}_t\right) - \frac{\text{Regret}(\mathcal{L}_{\mathcal{X}})}{T} \quad\geq\quad \min_{\mathbf{y}\in\mathcal{Y}}\max_{\mathbf{x}\in\mathcal{X}} \mathbf{x}^\top M\mathbf{y} - \frac{o(T)}{T}. \quad (2)$$

Combining these two statements gives $\min_{\mathbf{y}\in\mathcal{Y}}\max_{\mathbf{x}\in\mathcal{X}} \mathbf{x}^\top M\mathbf{y} \leq \max_{\mathbf{x}\in\mathcal{X}}\min_{\mathbf{y}\in\mathcal{Y}} \mathbf{x}^\top M\mathbf{y} + \frac{o(T)}{T}$. Of course, we can let $T \to \infty$ which immediately gives the desired inequality.

The previous example foreshadows a key result of this paper, which is that any no-regret learning algorithm can be converted into an approachability strategy. If we interpret Blackwell Approachability as a generalized form of Minimax Duality for vector-valued games then it may come as no surprise that regret-minimizing algorithms would provide a tool in establishing both game-theoretic results. However, in a certain sense regret-minimization is too heavy a hammer for proving Minimax Duality. For one, the above proof requires that we imagine a repeated version of the game, whereas scalar-valued game duality holds even for one-shot. Indeed, more standard proofs of von Neumann's result do not rely on repeated play. Blackwell Approachability, on the other hand, fundamentally involves repeated play, and in fact we shall show that regret-minimization is the perfectly-sized hammer, as it is *algorithmically equivalent* to approachability.

4. Equivalence of Approachability and Regret Minimization

4.1. Convex Cones and Conic Duality

We shall define some basic notions and then state some simple lemmas. Henceforth we use the notation $B_2(r)$ to refer to the ℓ_2-norm ball of radius r. The notation $\mathbf{x}' \oplus \mathbf{x}$ is the vector concatenation of $\mathbf{x}$ and $\mathbf{x}'$.

Definition 11 *A set $X \subset \mathbb{R}^d$ is a* cone *if it is closed under multiplication by nonnegative scalars, and X is a convex cone if it is also closed under element addition. Given any set $K \subset \mathbb{R}^d$, define the* conic hull $\texttt{cone}(K) := \{\alpha\mathbf{x} : \alpha \in \mathbb{R}_+, \mathbf{x} \in K\}$ *which is also a cone in $\mathbb{R}^d$. Also, given any convex cone $C \subset \mathbb{R}^d$, we can define the* polar cone *of C as*

$$C^0 := \{\boldsymbol{\theta} \in \mathbb{R}^d : \langle \boldsymbol{\theta}, \mathbf{x}\rangle \leq 0 \textit{ for all } \mathbf{x} \in C\}.$$

It is easily checked that if K is convex then $\texttt{cone}(K)$ is also convex. The following Lemma is folklore.

Lemma 12 *If C is a convex cone then (1) $(C^0)^0 = C$ and (2) supporting hyperplanes in C^0 correspond to points $\mathbf{x} \in C$, and vice versa. That is, given any supporting hyperplane H of C^0, H can be written exactly as $\{\boldsymbol{\theta} \in \mathbb{R}^d : \langle \boldsymbol{\theta}, \mathbf{x}\rangle = 0\}$ for some vector $\mathbf{x} \in C$ that is unique up to scaling.*

The distance to a cone can conveniently be measure via a "dual formulation," as we now show.

Lemma 13 *For every convex cone C in $\mathbb{R}^d$*

$$\texttt{dist}(\mathbf{x}, C) = \max_{\boldsymbol{\theta} \in C^0 \cap B_2(1)} \langle \boldsymbol{\theta}, \mathbf{x} \rangle \tag{3}$$

Proof We need two simple observations. Define $\pi_C(\mathbf{x})$ as the projection of $\mathbf{x}$ onto C. Then clearly, for any $\mathbf{x}$,

$$\texttt{dist}(\mathbf{x}, C) = \|\mathbf{x} - \pi_C(\mathbf{x})\| \tag{4}$$

$$\langle \mathbf{x} - \pi_C(\mathbf{x}), \mathbf{y} \rangle \leq 0 \ \forall \mathbf{y} \in C \text{ and hence } \mathbf{x} - \pi_C(\mathbf{x}) \in C^0 \tag{5}$$

$$\langle \mathbf{x} - \pi_C(\mathbf{x}), \pi_C(\mathbf{x}) \rangle = 0 \tag{6}$$

Given any $\boldsymbol{\theta} \in C^0$ with $\|\boldsymbol{\theta}\| \leq 1$, since $\pi_C(\mathbf{x}) \in C$ we have that

$$\langle \boldsymbol{\theta}, \mathbf{x} \rangle \leq \langle \boldsymbol{\theta}, \mathbf{x} - \pi_C(\mathbf{x}) \rangle \leq \|\boldsymbol{\theta}\| \|\mathbf{x} - \pi_C(\mathbf{x})\| \leq \|\mathbf{x} - \pi_C(\mathbf{x})\|,$$

which immediately implies that $\max_{\boldsymbol{\theta} \in C^0, \|\boldsymbol{\theta}\| \leq 1} \langle \boldsymbol{\theta}, \mathbf{x} \rangle \leq \texttt{dist}(\mathbf{x}, C)$. Furthermore, by selecting $\boldsymbol{\theta} = \frac{\mathbf{x} - \pi_C(\mathbf{x})}{\|\mathbf{x} - \pi_C(\mathbf{x})\|}$ which has norm one and, by (4), is in C^0, we see that

$$\max_{\boldsymbol{\theta} \in C^0, \|\boldsymbol{\theta}\| \leq 1} \langle \boldsymbol{\theta}, \mathbf{x} \rangle \geq \left\langle \frac{\mathbf{x} - \pi_C(\mathbf{x})}{\|\mathbf{x} - \pi_C(\mathbf{x})\|}, \mathbf{x} \right\rangle = \left\langle \frac{\mathbf{x} - \pi_C(\mathbf{x})}{\|\mathbf{x} - \pi_C(\mathbf{x})\|}, \mathbf{x} - \pi_C(\mathbf{x}) \right\rangle = \|\mathbf{x} - \pi_C(\mathbf{x})\|,$$

which implies that $\max_{\boldsymbol{\theta} \in C^0, \|\boldsymbol{\theta}\| \leq 1} \langle \boldsymbol{\theta}, \mathbf{x} \rangle \geq \texttt{dist}(\mathbf{x}, C)$ and hence we are done. ■

Our results require looking at convex cones rather than convex sets, hence we must consider the process of converting a set into a cone. In order to not lose information about the underlying set $\mathcal{K} \subset \mathbb{R}^d$, we shall embed the set into a higher dimension, and instead look at $\texttt{cone}(\{\kappa\} \times \mathcal{K}) \subset \mathbb{R}^{d+1}$, where $\kappa := \max_{\mathbf{x} \in \mathcal{K}} \|\mathbf{x}\|$ is the diameter of $\mathcal{K}$. We prove that this process of "lifting" and conifying does not perturb distances by more than a constant.

Lemma 14 *Consider a compact convex set $\mathcal{K} \subseteq \mathcal{H}$ in $\mathbb{R}^d$ and $\mathbf{x} \notin \mathcal{K}$. Let $\tilde{\mathbf{x}} := \kappa \oplus \mathbf{x}$ and $\tilde{\mathcal{K}} := \{\kappa\} \times \mathcal{K}$. Then we have*

$$\texttt{dist}(\tilde{\mathbf{x}}, \texttt{cone}(\tilde{\mathcal{K}})) \leq \texttt{dist}(\mathbf{x}, \mathcal{K}) \leq 2\texttt{dist}(\tilde{\mathbf{x}}, \texttt{cone}(\tilde{\mathcal{K}})) \tag{7}$$

Proof Since $\texttt{dist}(\tilde{\mathbf{x}}, \tilde{\mathcal{K}}) = \texttt{dist}(\mathbf{x}, \mathcal{K})$ and $\tilde{\mathcal{K}} \subset \texttt{cone}(\tilde{\mathcal{K}})$, the first inequality follows immediately.

For notational convenience let $\mathbf{w} = \pi_{\texttt{cone}(\tilde{\mathcal{K}})}(\mathbf{y})$ be the projection of $\mathbf{y}$ onto $\texttt{cone}(\tilde{\mathcal{K}})$ and $\mathbf{v} = \pi_{\tilde{\mathcal{K}}}(\mathbf{y})$ be the projection onto $\tilde{\mathcal{K}}$. Consider the plane determined by the three points $\tilde{\mathbf{x}}, \mathbf{w}, \mathbf{v}$. Notice that the triangle $\Delta(\tilde{\mathbf{x}}, \mathbf{w}, \mathbf{v})$ is similar to the triangle $\Delta(\mathbf{0}, \kappa \oplus \mathbf{0}, \mathbf{v})$, and hence by triangle similarity

$$\frac{\|\mathbf{v}\|}{\|\kappa \oplus \mathbf{0}\|} = \frac{\|\tilde{\mathbf{x}} - \mathbf{v}\|}{\|\tilde{\mathbf{x}} - \mathbf{w}\|} = \frac{\texttt{dist}(\tilde{\mathbf{x}}, \tilde{\mathcal{K}})}{\texttt{dist}(\tilde{\mathbf{x}}, \texttt{cone}(\tilde{\mathcal{K}}))}$$

For a visual aid, we provide a picture of this triangle similarity in Figure 1. Since $\mathbf{v} \in \tilde{\mathcal{K}}$ we have $\|\mathbf{v}\| \leq \|\tilde{\mathcal{K}}\| \leq 2\kappa$. In addition $\|\kappa \oplus \mathbf{0}\| = \kappa$ and the result follows. ■

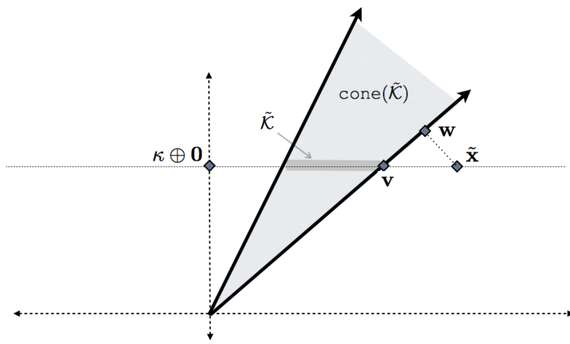

Figure 1: A geometric interpretation of the proof of Lemma 14.

4.2. Duality Theorems

In the previous sections we have presented two sequential decision problems, summarized in Figure 2. We now show that these two decision problems are *algorithmically equivalent*: any strategy (algorithm) that achieves approachability can be converted into an algorithm that achieves low-regret, and vice versa.

Blackwell Approachability Problem

Given a Blackwell instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S)$ and a valid half-space oracle $\mathcal{O} : H \mapsto \mathbf{x}_H \in \mathcal{X}$, construct an algorithm $\mathcal{A}$ so that, for any sequence $\mathbf{y}_1, \mathbf{y}_2, \ldots \in \mathcal{Y}$,

$$\texttt{dist}\left(\frac{1}{T}\sum_{t=1}^{T} \mathbf{u}(\mathbf{x}_t, \mathbf{y}_t), S\right) \to 0$$

where $\mathbf{x}_t \leftarrow \mathcal{A}(\mathbf{y}_1, \ldots, \mathbf{y}_{t-1})$.

Online Linear Optimization Problem

Given a compact convex set $\mathcal{K} \subset \mathbb{R}^d$, construct a learning algorithm $\mathcal{L}$ so that, for any sequence of loss vectors $\mathbf{f}_1, \mathbf{f}_2, \ldots \in \mathbb{R}^d$ we have vanishing regret, that is

$$\sum_{t=1}^{T} \langle \mathbf{f}_t, \mathbf{x}_t \rangle - \min_{\mathbf{x} \in K} \sum_{t=1}^{T} \langle \mathbf{f}_t, \mathbf{x} \rangle = o(T),$$

where $\mathbf{x}_t \leftarrow \mathcal{L}(\mathbf{f}_1, \ldots, \mathbf{f}_{t-1})$.

Figure 2: A summary of Blackwell Approachability and Online Linear Optimization

We present this equivalence as a pair of reductions. In Algorithm 1 we show how a learner, presented with a OLO problem characterized by a decision set $\mathcal{K}$ and an arriving sequence of loss vectors $\mathbf{f}_1, \mathbf{f}_2, \ldots$, can minimize regret with only oracle access to some approachability algorithm $\mathcal{A}$. In Algorithm 2 we show how a player, presented with a Blackwell instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S)$ and a valid halfspace oracle $\mathcal{O}$, can achieve approachability when only given oracle access to a no-regret OLO algorithm $\mathcal{L}$. For the remainder of the paper, for a given Blackwell instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S)$ and approachability algorithm

$\mathcal{A}$, $D(\mathcal{A}; \mathbf{y}_1, \ldots, \mathbf{y}_T)$ shall refer to the rate of approachability $\mathtt{dist}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{u}(\mathbf{x}_t, \mathbf{y}_t), S\right)$. We shall write $D_T(\mathcal{A})$ when the input sequence is clear. For the convex set $\mathcal{K}$, we shall let $\kappa := \max_{\mathbf{x}\in\mathcal{K}} \|\mathbf{x}\|$, the "norm" of the set $\mathcal{K}$.

Algorithm 1 Conversion of Approachability Alg. $\mathcal{A}$ to Online Linear Optimization Alg. $\mathcal{L}$

Input: compact convex decision set $\mathcal{K} \subset \mathbb{R}^d$
Input: sequence of cost functions $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_T \in B_2(1)$
Input: approachability oracle $\mathcal{A}$
Set: Blackwell instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S)$, where $\mathcal{X} := \mathcal{K}$, $\mathcal{Y} := B_2(1)$, $\mathbf{u}(\mathbf{x}, \mathbf{f}) = \frac{\langle \mathbf{f}, \mathbf{x}\rangle}{\kappa} \oplus -\mathbf{f}$, and $S := \mathtt{cone}(\{\kappa\} \times \mathcal{K})^0$
Construct: valid halfspace oracle $\mathcal{O}$ // Existence established in Lemma 15
for $t = 1, \ldots, T$ **do**
  Let: $\mathcal{L}(\mathbf{f}_1, \ldots, \mathbf{f}_{t-1}) := \mathcal{A}^{\mathcal{O}}_{\mathcal{X},\mathcal{Y},\mathbf{u},S}(\mathbf{f}_1, \ldots, \mathbf{f}_{t-1})$
  Receive: cost function $\mathbf{f}_t$
end for

In Algorithm 1 we require the construction of a valid halfspace oracle. In the lemma below we give one such oracle and prove that it is valid, but we note that this construction may not be the most efficient in general; any particular scenario may give rise to a simpler and faster construction.

Lemma 15 *There exists a valid halfspace oracle for the Blackwell instance in Algorithm 1.*

Proof Assume we have some halfspace H which contains $S = \mathtt{cone}(\{\kappa\} \times \mathcal{K})^0$. We can assume without loss of generality that H is tangent to S and, since S is a cone, H meets the origin; that is, $H = \{\boldsymbol{\theta} : \langle \boldsymbol{\theta}, \mathbf{z}_H\rangle \le 0\}$ for some $\mathbf{z}_H \in \mathbb{R}^d$. Furthermore, $H \supset \mathtt{cone}(\{\kappa\} \times \mathcal{K})^0$ implies that $\mathbf{z}_H \in (\mathtt{cone}(\{\kappa\} \times \mathcal{K})^0)^0 = \mathtt{cone}(\{\kappa\} \times \mathcal{K})$. Equivalently, $\mathbf{z}_H = \alpha(\kappa \oplus \mathbf{x}_H)$ for some $\mathbf{x}_H \in \mathcal{K}$ and some $\alpha > 0$. With this in mind, we construct our oracle by setting $\mathbf{x}_H \leftarrow \mathcal{O}(H)$.

It remains to prove that this halfspace oracle is valid. We compute $\langle \mathbf{u}(\mathbf{x}_H, \mathbf{f}), \mathbf{z}_H\rangle$:

$$\langle \mathbf{u}(\mathbf{x}_H, \mathbf{f}), \mathbf{z}_H\rangle = \langle \kappa^{-1}\langle \mathbf{f}, \mathbf{x}_H\rangle \oplus -\mathbf{f}, \alpha\kappa \oplus \alpha\mathbf{x}_H\rangle = \alpha\langle \mathbf{f}, \mathbf{x}_H\rangle + \langle -\mathbf{f}, \alpha\mathbf{x}_H\rangle = 0.$$

By definition, $\langle \mathbf{u}(\mathbf{x}_H, \mathbf{f}), \mathbf{z}_H\rangle \le 0$ implies that $\mathbf{u}(\mathbf{x}_H, \mathbf{f}) \in H$ for any $\mathbf{f}$ and we are done. ■

Theorem 16 *The reduction defined in Algorithm 1, for any input algorithm $\mathcal{A}$, produces an OLO algorithm $\mathcal{L}$ such that $\frac{\mathrm{Regret}(\mathcal{L})}{T} \le 2\kappa D_T(\mathcal{A})$.*

Proof Applying Lemmas 13 and 12 to the definition of $D_T(\mathcal{A})$ gives

$$D_T(\mathcal{A}) \equiv \mathtt{dist}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{u}(\mathbf{x}_t, \mathbf{f}_t), S\right) = \max_{\mathbf{w}\in\mathtt{cone}(\kappa\oplus\mathcal{K})\cap B_2^d(1)} \left\langle \frac{1}{T}\sum_{t=1}^{T}\mathbf{u}(\mathbf{x}_t, \mathbf{f}_t), \mathbf{w}\right\rangle \qquad (8)$$

Notice that, in this optimization, we can assume w.l.o.g. that $\|\mathbf{w}\| = 1$, or $\mathbf{w} = \mathbf{0}$. In the former case we can write $\mathbf{w} = \frac{\kappa \oplus \mathbf{x}}{\|\kappa \oplus \mathbf{x}\|}$ for some $\mathbf{x} \in \mathcal{K}$, and we drop the latter case to obtain the inequality

$$\begin{aligned} D_T(\mathcal{A}) &\geq \max_{\mathbf{x} \in \mathcal{K}} \left\langle \frac{1}{T} \sum_{t=1}^{T} \mathbf{u}(\mathbf{x}_t, \mathbf{f}_t), \frac{\kappa \oplus \mathbf{x}}{\|\kappa \oplus \mathbf{x}\|} \right\rangle \\ &= \frac{1}{T} \max_{\mathbf{x} \in \mathcal{K}} \frac{\left(\sum_{t=1}^{T} \langle \mathbf{f}_t, \mathbf{x}_t \rangle - \sum_{t=1}^{T} \langle \mathbf{f}_t, \mathbf{x} \rangle \right)}{\|\kappa \oplus \mathbf{x}\|} \\ &\geq \frac{\frac{1}{T} \left(\sum_{t=1}^{T} \langle \mathbf{f}_t, \mathbf{x}_t \rangle - \sum_{t=1}^{T} \langle \mathbf{f}_t, \mathbf{x}^* \rangle \right)}{\|\kappa \oplus \mathbf{x}^*\|} \geq \frac{\frac{1}{T}\mathrm{Regret}_T(\mathcal{A})}{2\kappa}, \end{aligned}$$

where we set $\mathbf{x}^* := \arg\min_{\mathbf{x} \in \mathcal{K}} \sum_{t=1}^{T} \langle \mathbf{f}_t, \mathbf{x} \rangle$. ∎

We turn our attention to the second reduction.

Algorithm 2 Conversion of Online Linear Optimization Alg. $\mathcal{L}$ to Approachability Alg. $\mathcal{A}$

1: Input: Blackwell instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S)$, with S a cone; and a valid halfspace oracle $\mathcal{O}$
2: Input: Online Linear Optimization oracle $\mathcal{L}$
3: Set: $\mathcal{K} = S^0 \cap B_2(1)$
4: **for** $t = 1, \ldots, T$ **do**
5: Query $\mathcal{L}$: $\boldsymbol{\theta}_t \leftarrow \mathcal{L}_{\mathcal{K}}(\mathbf{f}_1, \ldots, \mathbf{f}_{t-1})$, where $\mathbf{f}_s \leftarrow -\mathbf{u}(\mathbf{x}_s, \mathbf{y}_s)$
6: Query $\mathcal{O}$: $\mathbf{x}_t \leftarrow \mathcal{O}(H_{\boldsymbol{\theta}_t})$ where $H_{\boldsymbol{\theta}_t} := \{\mathbf{z} : \langle \boldsymbol{\theta}_t, \mathbf{z} \rangle \leq 0\}$
7: Let: $\mathcal{A}(\mathbf{y}_1, \ldots, \mathbf{y}_{t-1}) := \mathbf{x}_t$
8: Receive: $\mathbf{y}_t \in \mathcal{Y}$
9: **end for**

We now prove a similar rate for reverse direction. Here we assume that S is a cone, but we relax this restriction next.

Theorem 17 *The reduction in Algorithm 2, when S is a cone, leads to a rate of approachability of algorithm $\mathcal{A}$ of $D_T(\mathcal{A}; \mathbf{y}_{1:T}) \leq \frac{\mathrm{Regret}(\mathcal{L}_{\mathcal{K}}; \mathbf{f}_{1:T})}{T}$.*

Proof We state precisely the halfspace oracle guarantee from line 6. We know that $\mathbf{u}(\mathbf{x}_t, \mathbf{y}) \in H_{\boldsymbol{\theta}_t}$ or equivalently $\langle \boldsymbol{\theta}_t, \mathbf{u}(\mathbf{x}_t, \mathbf{y}) \rangle \leq 0$ for any $\mathbf{y} \in \mathcal{Y}$. In particular, since $\mathbf{u}(\mathbf{x}_t, \mathbf{y}_t) = -\mathbf{f}_t$, we have $\langle \boldsymbol{\theta}_t, \mathbf{f}_t \rangle \geq 0$. We bound $D_T(\mathcal{A})$ by applying Lemma 13 to obtain:

$$\begin{aligned} D_T(\mathcal{A}) = \mathtt{dist}\left(\frac{1}{T} \sum_{t=1}^{T} \mathbf{u}(\mathbf{x}_t, \mathbf{y}_t), S \right) &= \max_{\boldsymbol{\theta} \in \mathcal{K}} \left\langle \frac{1}{T} \sum_{t=1}^{T} \mathbf{u}(\mathbf{x}_t, \mathbf{y}_t), \boldsymbol{\theta} \right\rangle = \frac{1}{T} \max_{\boldsymbol{\theta} \in \mathcal{K}} \left(-\sum_{t=1}^{T} \langle \mathbf{f}_t, \boldsymbol{\theta} \rangle \right) \\ \leq \frac{1}{T} \left(\sum_{t=1}^{T} \langle \mathbf{f}_t, \boldsymbol{\theta}_t \rangle - \min_{\boldsymbol{\theta} \in \mathcal{K}} \sum_{t=1}^{T} \langle \mathbf{f}_t, \boldsymbol{\theta} \rangle \right) &= \frac{1}{T} \mathrm{Regret}_T(\mathcal{A}) \end{aligned} \tag{9}$$

where the inequality follows by the halfspace oracle guarantee. ∎

For a Blackwell instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S)$, even when S is not a cone we can still use Algorithm 2 by *lifting* S: apply Algorithm 2 to the instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}'(\cdot,\cdot), S')$, where $S' := \texttt{cone}(\{\kappa\} \times S)$ and $\mathbf{u}'(\mathbf{x}, \mathbf{y}) := \kappa \oplus \mathbf{u}(\mathbf{x}, \mathbf{y})$.

Corollary 18 *Given a Blackwell instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S)$ with compact S, and let its lifted instance be $(\mathcal{X}, \mathcal{Y}, \mathbf{u}'(\cdot,\cdot), S')$ as described above. Then*

$$\texttt{dist}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{u}(\mathbf{x}_t, \mathbf{y}_t), S\right) \leq 2 \cdot \texttt{dist}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{u}'(\mathbf{x}_t, \mathbf{y}_t), S'\right) \leq \frac{2}{T}\text{Regret}_T(\mathcal{A})$$

Proof Apply Lemma 14 to Theorem 17. ■

We include the compactness assumption only because Lemma 14 requires it yet it is not necessary; the size of S does not enter into the bound. For any Blackwell instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S)$ with non-compact S, we may always consider a functionally equivalent instance $(\mathcal{X}, \mathcal{Y}, \mathbf{u}(\cdot,\cdot), S_0)$, where $S_0 \subset S$ is compact. Letting $U := \{\mathbf{u}(\mathbf{x}, \mathbf{y}) : \mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}\}$, which is compact, we may simply let S_0 be the convex hull of all projections of points in U onto S. Hence $\texttt{dist}(\mathbf{z}, S) = \texttt{dist}(\mathbf{z}, S_0)$ for all $\mathbf{z} \in U$.

5. Efficient Calibration via Approachability and OLO

Imagine a sequence of binary outcomes, say ‘rain’ or ‘shine’ on a given day, and imagine a forecaster, say the weatherman, that wants to predict the probability of this outcome on each day. A natural question to ask is, on the days when the weatherman actually predicts “30% chance of rain”, does it actually rain (roughly) 30% of the time? This exactly the problem of *calibrated forecasting* which we now discuss.

There have been a range of definitions of calibration given throughout the literature, some equivalent and some not, but from a computational viewpoint there are significant differences. We thus give a clean definition of calibration, first introduced by Foster (1999), which is convenient to asses computationally.

We let $y_1, y_2, \ldots \in \{0, 1\}$ be a sequence of outcomes, and $p_1, p_2, \ldots \in [0, 1]$ a sequence of probability predictions by a forecaster. We define for every T and every probability interval $[p - \varepsilon/2, p + \varepsilon/2)$ for $p \in [0, 1]$ and $\varepsilon > 0$, the quantities

$$n_T(p, \varepsilon) := \sum_{t=1}^{T}\mathbb{I}[p_t \in [p - \varepsilon/2, p + \varepsilon/2)], \qquad \rho_T(p, \varepsilon) := \frac{\sum_{t=1}^{T} y_t \mathbb{I}[p_t \in [p - \varepsilon/2, p + \varepsilon/2)]}{n_T(p, \varepsilon)}.$$

The quantity $\rho_T(p, \varepsilon)$ should be interpreted as the empirical frequency of $y_t = 1$, up to round T, on only those rounds where the forecaster's prediction was within $\varepsilon/2$ of p. The goal of calibration, of course, is to have this empirical frequency $\rho_T(p, \varepsilon)$ be close to the estimated frequency p in the limit. The standard definition of a calibrated forecaster is one that satisfies

$$\text{for all } p \in [0, 1], \varepsilon > 0 : \quad \limsup_{T\to\infty} |\rho_T(p, \varepsilon) - p| \leq O(\varepsilon) \quad \text{unless} \quad n_T(p, \varepsilon) = o(T). \tag{10}$$

Requiring that $n_T(p,\varepsilon)$ does not grow too slowly is an important condition, as we can not expect the forecaster to be calibrated in regions on which he predicts only a small number of times. On the other hand, this case-sensitive condition is somewhat awkward, and we instead use the following equivalent notion.

Definition 19 *Let the (ℓ_1,ε)-calibration rate for forecaster $\mathcal{A}$ be*

$$C_T^\varepsilon(\mathcal{A}) = \max\left\{0, \sum_{i=0}^{\lfloor \varepsilon^{-1} \rfloor} \frac{n_T(i\varepsilon,\varepsilon)}{T} \left| i\varepsilon - \rho_T(i\varepsilon,\varepsilon) \right| - \frac{\varepsilon}{2} \right\}.$$

*We say that a forecaster is (ℓ_1,ε)-*calibrated *if $C_T^\varepsilon(\mathcal{A}) = o(1)$.*

The definition of asymptotic calibration considers the "total error" over an ε-grid, and it adjusts the normalization for each term to $\frac{1}{T}$. The benefit here is that we can ignore intervals in this grid for which $n_T(p,\varepsilon) = o(T)$. In addition, we subtract the constant $\varepsilon/2$ which is an artifact of the discretization by ε; this is the smallest constant which allows for $\limsup_{T\to\infty} C_T^\varepsilon(\mathcal{A}) \le 0$. A standard reduction in the literature (see e.g. (Cesa-Bianchi and Lugosi, 2006)) shows that a fully-calibrated algorithm (i.e. one satisfying (10)) can be constructed from and (ℓ_1,ε)-calibrated algorithm. Henceforth we only consider the (ℓ_1,ε) condition.

As our goal is to minimize the calibration score C_T^ε, we can interpret this value instead as a distance to the ℓ_1-norm ball. Define the *calibration vector* $\mathbf{c}_T \in \mathbb{R}^{\lfloor \varepsilon^{-1} \rfloor}$ at time T as: $\mathbf{c}_T(i) = \frac{n_T(i\varepsilon,\varepsilon)}{T}(i\varepsilon - \rho_T(i\varepsilon,\varepsilon))$.

Claim 1 *Whenever $\mathbf{c}_T \notin B_1(\varepsilon/2)$, we have*

$$C_T^\varepsilon = \mathtt{dist}_1(\mathbf{c}_T, B_1(\varepsilon/2)).$$

Proof Notice that for any $\mathbf{x}$: $\mathtt{dist}_1(\mathbf{x}, B_1(\varepsilon/2)) := \min_{\mathbf{y}:\|\mathbf{y}\|_1 \le \varepsilon/2} \|\mathbf{x}-\mathbf{y}\|_1 = \max\{0, -\varepsilon/2 + \|\mathbf{x}\|_1\}$. The second equality follows by noting that an optimally chosen $\mathbf{y}$ will lie in the same quadrant as $\mathbf{x}$. When we set $\mathbf{x} = \mathbf{c}_T$, it is clear that $\|\mathbf{c}_T\|_1 > \varepsilon/2$ given our assumption that $\mathbf{c}_T \notin B_1(\varepsilon/2)$. ∎

The utility of this claim shall be to convert the problem of (ℓ_1,ε)-calibration to a problem of approachability; that is, can we approach the set $B_1(\varepsilon/2)$ for a particular vector-valued game? In the following section we describe this construction in detail.

5.1. Existence of Calibrated Forecaster via Blackwell Approachability

A surprising fact is that it is possible to achieve calibration even when the outcome sequence $\{y_t\}$ is chosen by an adversary, although this requires a randomized strategy of the forecaster. Algorithms for calibrated forecasting under adversarial conditions have been given in Foster and Vohra (1998), Fudenberg and Levine (1999), and Hart and Mas-Colell (2000).

Interestingly, the calibration problem was reduced to Blackwell's Approachability Theorem in a short paper by Foster (1999). Foster's reduction uses Blackwell's original theorem, proving that a given set is halfspace-satisfiable, in particular by providing a construction for each such halfspace. Here we provide a reduction to Blackwell Approachability using the response-satisfiability condition – that is by using Theorem 6 – which is both significantly easier and more intuitive than Foster's construction[2]. We also show, using the reduction to Online Linear Optimization from the previous section, how to achieve the most efficient known algorithm for calibration by taking advantage of the Online Gradient Descent algorithm of Zinkevich (2003), using the results of Section 4.

We now describe the construction that allows us to reduce calibration to approachability. For any $\varepsilon > 0$ we will show how to construct an (ℓ_1, ε)-calibrated forecaster. Notice that from here, it is straightforward to produce a well-calibrated forecaster (Foster and Vohra, 1998). For simplicity, assume $\varepsilon = 1/m$ for some positive integer m. On each round t, a forecaster will now randomly predict a probability $p_t \in \{0/m, 1/m, 2/m, \ldots, (m-1)/m, 1\}$, according to the distribution $\mathbf{w}_t$, that is $\Pr(p_t = i/m) = \mathbf{w}_t(i)$. We now define a vector-valued game. Let the player choose $\mathbf{w}_t \in \mathcal{X} := \Delta_{m+1}$, and the adversary choose $y_t \in \mathcal{Y} := [0,1]$, and the payoff vector will be

$$\mathbf{u}(\mathbf{w}_t, y_t) := \left\langle \mathbf{w}_t(0)\left(y_t - \frac{0}{m}\right), \mathbf{w}_t(1)\left(y_t - \frac{1}{m}\right), \ldots, \mathbf{w}_t(m)(y_t - 1)\right\rangle \tag{11}$$

Lemma 20 *Consider the vector-valued game described above and let $S := B_1(\varepsilon/2)$. If we have a strategy for choosing $\mathbf{w}_t$ that guarantees approachability of S, that is $\frac{1}{T}\sum_{t=1}^T \mathbf{u}(\mathbf{w}_t, y_t) \to S$, then a randomized forecaster that selects p_t according to $\mathbf{w}_t$ is (ℓ_1, ε)-calibrated with high probability.*

The proof of this lemma is straightforward, and is similar to the construction in Foster (1999). The key fact is that $\frac{1}{T}\sum_{t=1}^T \mathbf{u}(\mathbf{w}_t, y_t) = \mathbb{E}[\mathbf{c}_T]$, where the expectation is taken over the algorithms draws of every p_t according to the distribution $\mathbf{w}_t$. Since each p_t is drawn independently, by standard concentration arguments we can see that if $\frac{1}{T}\sum_{t=1}^T \mathbf{u}(\mathbf{w}_t, y_t)$ is close to the $\ell_1$ball of radius $\varepsilon/2$, then the (ℓ_1, ε)-calibration vector is close to the $\varepsilon/2$ ball with high probability.

We can now apply Theorem 6 to prove the existence of a calibrated forecaster.

Theorem 21 *For the vector-valued game defined in* (11)*, the set $B_1(\varepsilon/2)$ is response-satisfiable and, hence, approachable.*

Proof To show response-satisfiability, we need only show that, for every strategy $y \in [0,1]$ played by the adversary, there is a strategy $\mathbf{w} \in \Delta_m$ for which $\mathbf{u}(\mathbf{w}, y) \in S$. This can be achieved by simply setting i so as to minimize $|i\varepsilon - y|$, which can always be made smaller than $\varepsilon/2$. We then choose our distribution $\mathbf{w} \in \Delta_{m+1}$ to be a point mass on i, that is we set $w(i) = 1$ and $w(j) = 0$ for all $j \neq i$. Then $\mathbf{u}(\mathbf{w}, y)$ is identically 0 everywhere except the ith coordinate, which has the value $y - i/m$. By construction, $y - i/m \in [-1/m, 1/m]$, and we are done. ■

2. A similar existence proof was discovered concurrently by Mannor and Stoltz (2009)

5.2. Efficient Algorithm for Calibration via Online Linear Optimization

We now show how the results in the previous Section lead to the first efficient algorithm for calibrated forecasting. The previous theorem provides a natural existence proof for Calibration, but it does not immediately provide us with a simple and efficient algorithm. We proceed according to the reduction outlined in the previous section to prove:

Theorem 22 *There exists a (ℓ_1, ε)-calibration algorithm that runs in time $O(\log \frac{1}{\varepsilon})$ per iteration and satisfies $C_T^\varepsilon = O\left(\frac{1}{\sqrt{\varepsilon T}}\right)$*

The reduction developed in Theorem 17 has some flexibility, and we shall modify it for the purposes of this problem. The objects we shall need, as well as the required conditions, are as follows:

1. A convex set $\mathcal{K}$
2. An efficient algorithm $\mathcal{A}$ which, for any sequence $\mathbf{f}_1, \mathbf{f}_2, \ldots$, can select a sequence of points $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots \in \mathcal{K}$ with the guarantee that $\sum_{t=1}^T \langle \mathbf{f}_t, \boldsymbol{\theta}_t \rangle - \min_{\boldsymbol{\theta} \in \mathcal{K}} \sum_{t=1}^T \langle \mathbf{f}_t, \boldsymbol{\theta} \rangle = o(T)$. For the reduction, we shall set $\mathbf{f}_t \leftarrow -\mathbf{u}(\mathbf{w}_t, y_t)$.
3. An efficient oracle that can select a particular $\mathbf{w}_t \in \mathcal{X}$ for each $\boldsymbol{\theta}_t \in \mathcal{K}$ with the guarantee that

$$\texttt{dist}\left(\frac{1}{T}\sum_{t=1}^T \mathbf{u}(\mathbf{w}_t, y_t), S\right) \leq \frac{1}{T}\left(\sum_{t=1}^T \langle -\mathbf{u}(\mathbf{w}_t, y_t), \boldsymbol{\theta}_t \rangle - \min_{\boldsymbol{\theta} \in \mathcal{K}} \sum_{t=1}^T \langle -\mathbf{u}(\mathbf{w}_t, y_t), \boldsymbol{\theta} \rangle\right) \tag{12}$$

   where the function $\texttt{dist}()$ can be with respect to any norm.

The Setup Let $\mathcal{K} = B_\infty(1) = \{\boldsymbol{\theta} \in \mathbb{R}^d : \|\boldsymbol{\theta}\|_\infty \leq 1\}$ be the unit cube. This is an appropriate choice because we can write $\texttt{dist}_1(\mathbf{x}, B_1(\varepsilon/2))$ for $\mathbf{x} \notin B_1(\varepsilon/2))$ as

$$\texttt{dist}_1(\mathbf{x}, B_1(\varepsilon/2)) := \min_{\mathbf{y}:\|\mathbf{y}\|_1 \leq \varepsilon/2} \|\mathbf{x} - \mathbf{y}\|_1 = -\varepsilon/2 + \|\mathbf{x}\|_1 = -\varepsilon/2 - \min_{\boldsymbol{\theta}:\|\boldsymbol{\theta}\|_\infty \leq 1} \langle -\mathbf{x}, \boldsymbol{\theta} \rangle; \tag{13}$$

The former equality was proved in Claim 1. Furthermore, we shall construct our oracle mapping $\boldsymbol{\theta} \mapsto \mathbf{w}$ with the following guarantee: $\langle \mathbf{u}(\mathbf{w}, y), \boldsymbol{\theta} \rangle \leq \varepsilon/2$ for any y. Using this guarantee, and if we plug in $\mathbf{x} = \frac{1}{T}\sum_{t=1}^T \mathbf{u}(\mathbf{w}_t, y_t)$ (13), we arrive at:

$$\begin{aligned}
\texttt{dist}_1\left(\frac{\sum_{t=1}^T \mathbf{u}(\mathbf{w}_t, y_t)}{T}, B_1(\varepsilon/2)\right) &= -\varepsilon/2 - \min_{\boldsymbol{\theta}:\|\boldsymbol{\theta}\|_\infty \leq 1} \left\langle \frac{-\sum_{t=1}^T \mathbf{u}(\mathbf{w}_t, y_t)}{T}, \boldsymbol{\theta} \right\rangle \\
&\leq \frac{1}{T}\left(\sum_{t=1}^T \langle -\mathbf{u}(\mathbf{w}_t, y_t), \boldsymbol{\theta}_t \rangle - \min_{\boldsymbol{\theta} \in \mathcal{K}} \sum_{t=1}^T \langle -\mathbf{u}(\mathbf{w}_t, y_t), \boldsymbol{\theta} \rangle\right)
\end{aligned}$$

This is precisely the necessary guarantee (12).

Algorithm 3 Efficient Oracle mapping $\mathcal{O} : \mathbf{w} \mapsto \boldsymbol{\theta}$

Input: $\boldsymbol{\theta}$ such that $\|\boldsymbol{\theta}\|_\infty \leq 1$
if $\boldsymbol{\theta}(0) \leq 0$ **then**
  $\mathbf{w} \leftarrow \delta_0$ // That is, choose $\mathbf{w}$ to place all weight on the 0th coordinate
else if $\theta(m) \geq 0$ **then**
  $\mathbf{w} \leftarrow \delta_m$ // That is, choose $\mathbf{w}$ to place all weight on the last coordinate
else
  Binary search $\boldsymbol{\theta}$ to find coordinate i such that $\boldsymbol{\theta}(i) > 0$ and $\boldsymbol{\theta}(i+1) \leq 0$
  $\mathbf{w} \leftarrow \frac{\boldsymbol{\theta}(i)^{-1}}{\boldsymbol{\theta}(i)^{-1} - \boldsymbol{\theta}(i+1)^{-1}} \delta_i + \frac{-\boldsymbol{\theta}(i+1)^{-1}}{\boldsymbol{\theta}(i)^{-1} - \boldsymbol{\theta}(i+1)^{-1}} \delta_{i+1}$
end if
Return $\mathbf{w}$

Constructing the Oracle We now turn our attention to designing the required oracle in an *efficient* manner. In particular, given any $\boldsymbol{\theta}$ with $\|\boldsymbol{\theta}\|_\infty \leq 1$ we must construct $\mathbf{w} \in \Delta_{m+1}$ so that $\langle \ell(\mathbf{w}, y), \boldsymbol{\theta} \rangle \leq \varepsilon/2$ for any y. The details of this oracle are given in Algorithm 3. It is straightforward why, in the final **else** condition, there must be such a pair of coordinates $i, i+1$ satisfying the condition. We need not be concerned with the case that $\boldsymbol{\theta}(i+1) = 0$, where we can simply define $\frac{0}{\infty} = 0$ and $\frac{\infty}{\infty} = 1$ leading to $\mathbf{w} \leftarrow \delta_{i+1}$. It is also clear that, with the binary search, this algorithm requires at most $O(\log m) = O(\log 1/\varepsilon)$ computation.

In order to prove that this construction is valid we need to check the condition that, for any $y \in \{0, 1\}$, $\langle \mathbf{u}(\mathbf{w}, y), \boldsymbol{\theta} \rangle \leq \varepsilon/2$; or more precisely, $\sum_{i=1}^m \boldsymbol{\theta}(i)\mathbf{w}(i)\left(y - \frac{i}{m}\right) \leq \varepsilon/2$. Recalling that $m = 1/\varepsilon$, this is trivially checked for the case when $\boldsymbol{\theta}(1) \leq 0$ or $\boldsymbol{\theta}(m) \geq 0$. Otherwise, we have

$$\begin{aligned}\langle \mathbf{u}(\mathbf{w}, y), \boldsymbol{\theta} \rangle &= \frac{\boldsymbol{\theta}(i) \cdot \boldsymbol{\theta}(i)^{-1}}{\boldsymbol{\theta}(i)^{-1} - \boldsymbol{\theta}(i+1)^{-1}} \left(y - \frac{i}{m}\right) + \frac{\boldsymbol{\theta}(i+1) \cdot (-\boldsymbol{\theta}(i+1)^{-1})}{\boldsymbol{\theta}(i)^{-1} - \boldsymbol{\theta}(i+1)^{-1}} \left(y - \frac{i+1}{m}\right) \\ &= \frac{1}{\boldsymbol{\theta}(i)^{-1} - \boldsymbol{\theta}(i+1)^{-1}} \frac{1}{m} \leq \frac{\max(|\boldsymbol{\theta}(i)|, |\boldsymbol{\theta}(i+1)|)}{2} \varepsilon \leq \frac{\varepsilon}{2}\end{aligned}$$

The Learning Algorithm The final piece is to construct an efficient learning algorithm which leads to vanishing regret. That is, we need to construct a sequence of $\boldsymbol{\theta}_t$'s in the unit cube (denoted $B_\infty(1)$) so that

$$\sum_{t=1}^T \langle \mathbf{u}_t, \boldsymbol{\theta}_t \rangle - \min_{\boldsymbol{\theta} \in B_\infty(1)} \sum_{t=1}^T \langle \mathbf{u}_t, \boldsymbol{\theta} \rangle = o(T),$$

where $\mathbf{u}_t := \mathbf{u}(\mathbf{w}_t, y_t)$. There are a range of possible no-regret algorithms available, but we use the one given by Zinkevich known commonly as Online Gradient Descent (Zinkevich, 2003). The details are given in Algorithm 4. This algorithm can indeed be implemented efficiently, requiring only $O(1)$ computation on each round and $O(\min\{m, T\})$ memory. The main advantage is that the vectors $\mathbf{u}_t$ are generated via our oracle above, and these vectors are *sparse*, having only at most two nonzero coordinates. Hence, the Gradient Descent Step requires only $O(1)$ computation. In addition, the Projection Step can also be performed in an efficient manner. Since we assume that $\boldsymbol{\theta}_t \in B_\infty(1)$, the updated point $\boldsymbol{\theta}'_{t+1}$ can

Algorithm 4 Online Gradient Descent

Input: convex set $\mathcal{K} \subset \mathbb{R}^d$
Initialize: $\boldsymbol{\theta}_1 = \mathbf{0}$
Set Parameter: $\eta = O(T^{-1/2})$
for $t = 1, \dots, T$ **do**
  Receive $\mathbf{u}_t$
  $\boldsymbol{\theta}'_{t+1} \leftarrow \boldsymbol{\theta}_t - \eta \mathbf{u}_t$ // Gradient Descent Step
  $\boldsymbol{\theta}_{t+1} \leftarrow \text{Project}_2(\boldsymbol{\theta}'_{t+1}, \mathcal{K})$ // L2 Projection Step
end for

violate at most two of the ℓ_∞ constraints of the ball $B_\infty(1)$. An ℓ_2 projection onto the cube requires simply rounding the violated coordinates into $[-1, 1]$. The number of non-zero elements in $\boldsymbol{\theta}$ can increase by at most two every iteration, and storing $\boldsymbol{\theta}$ is the only state that online gradient descent needs to store, hence the algorithm can be implemented with $O(\min\{T, m\})$ memory. We thus arrive at an efficient no-regret algorithm for choosing $\boldsymbol{\theta}_t$.

Putting it all Together We can now fully specify our calibration algorithm given the subroutines defined above. The precise description is in Algorithm 5, which makes queries to Algorithms 3 and 4.

Algorithm 5 Efficient Algorithm for Asymptotic Calibration

Input: $\varepsilon = 1/m$ for some natural number m
Initialize: $\boldsymbol{\theta}_1 = \mathbf{0}$, $\mathbf{w}_1 \in \Delta_{m+1}$ arbitrarily
for $t = 1, \dots, T$ **do**
  Sample $i_t \sim \mathbf{w}_t$, predict $p_t = \frac{i_t}{m}$, observe $y_t \in \{0, 1\}$
  Set $\mathbf{u}_t := \mathbf{u}(\mathbf{w}_t, y_t)$ // Vector-valued game defined in (11)
  Query learning algorithm: $\boldsymbol{\theta}_{t+1} \leftarrow \text{Update}(\boldsymbol{\theta}_t | \mathbf{u}_t)$ // Subroutine from Algorithm 4
  Query halfspace oracle: $\mathbf{w}_{t+1} \leftarrow \mathcal{O}(\boldsymbol{\theta}_{t+1})$ // Subroutine from Algorithm 3
end for

Proof [of Theorem 22] Here we have bounded the distance directly by the regret, using equation (12), which tells us that the calibration rate is bounded by the regret of the online learning algorithm. Online Gradient Descent guarantees the regret to be no more than $DG\sqrt{T}$, where D is the ℓ_2 diameter of the set, and G is the ℓ_2-norm of the largest cost vector. For the ball $B_\infty(1)$, the diameter $D = \sqrt{\frac{1}{\varepsilon}}$, and we can bound the norm of our loss vectors by $G = \sqrt{2}$. Hence:

$$C_T^\varepsilon \quad = \quad \texttt{dist}(c_T, B_1(\varepsilon/2)) \quad \leq \quad \frac{\text{Regret}_T}{T} \quad \leq \quad \frac{GD}{\sqrt{T}} \quad = \quad O\left(\frac{1}{\sqrt{\varepsilon T}}\right) \tag{14}$$

■

References

J. Abernethy, A. Agarwal, P.L. Bartlett, and A. Rakhlin. A stochastic view of optimal regret through minimax duality. In *Proceedings of the 22nd Annual Conference on Learning Theory*, 2009.

D. Blackwell. Controlled random walks. In *Proceedings of the International Congress of Mathematicians*, volume 3, pages 336–338, 1954.

D. Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):18, 1956.

Nicolo Cesa-Bianchi and Gabor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006. ISBN 0521841089, 9780521841085.

A. Dawid. The well-calibrated Bayesian. *Journal of the American Statistical Association*, 77:605–613, 1982.

E. Even-Dar, R. Kleinberg, S. Mannor, and Y. Mansour. Online learning for global cost functions. In *22nd Annual Conference on Learning Theory (COLT)*, 2009.

D. P Foster. A proof of calibration via blackwell's approachability theorem. *Games and Economic Behavior*, 29(1-2):7378, 1999.

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

Y. Freund and R. Schapire. A desicion-theoretic generalization of on-line learning and an application to boosting. In *Computational learning theory*, pages 23–37. Springer, 1995.

D. Fudenberg and D. K Levine. An easier way to calibrate* 1. *Games and economic behavior*, 29(1-2):131137, 1999.

J. Hannan. Approximation to Bayes risk in repeated play. *Contributions to the Theory of Games*, 3:97–139, 1957.

S. Hart and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):11271150, 2000.

E. Hazan, A. Agarwal, and S. Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2):169–192, 2007. ISSN 0885-6125.

Elad Hazan. The convex optimization approach to regret minimization. In *To appear in Optimization for Machine Learning*. MIT Press, 2010.

S. Mannor and N. Shimkin. Regret minimization in repeated matrix games with variable stage duration. *Games and Economic Behavior*, 63(1):227–258, 2008. ISSN 0899-8256.

Shie Mannor and Gilles Stoltz. A Geometric Proof of Calibration. *arXiv*, Dec 2009. URL `http://arxiv.org/abs/0912.3604`.

J. Von Neumann, O. Morgenstern, H. W Kuhn, and A. Rubinstein. *Theory of games and economic behavior*. Princeton university press Princeton, NJ, 1947.

V. Perchet. Calibration and internal no-regret with random signals. In *Proceedings of the 20th international conference on Algorithmic learning theory*, pages 68–82. Springer-Verlag, 2009. ISBN 3642044131.

A. Rakhlin, K. Sridharan, and A. Tewari. Online Learning: Beyond Regret. *Arxiv preprint arXiv:1011.3168*, 2010.

A. Sandroni, R. Smorodinsky, and R.V. Vohra. Calibration with many checking rules. *Mathematics of Operations Research*, 28(1):141–153, 2003. ISSN 0364-765X.

S. Shalev-Shwartz and Y. Singer. Convex repeated games and Fenchel duality. *Advances in Neural Information Processing Systems*, 19:1265, 2007. ISSN 1049-5258.

M. Sion. On general minimax theorems. *Pacific J. Math*, 8(1):171–176, 1958.

M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *International Conference on Machine Learning*, volume 20, page 928, 2003.