Journal of Machine Learning Research 26 (2025) 1-75 Submitted 11/24; Revised 4/25; Published 7/25

Imprecise Multi-Armed Bandits: Representing Irreducible Uncertainty as a Zero-Sum Game

Vanessa Kosoy VANESSA@ALTER.ORG.IL
Faculty of Mathematics
Technion - Israel Institute of Technology
Haifa 3200003, Israel

Association for Long Term Existence and Resilience
Rehovot, Israel

Editor: Kevin Jamieson

Abstract

We introduce a novel multi-armed bandit framework, where each arm is associated with a fixed unknown *credal set* over the space of outcomes (which can be richer than just the reward). The arm-to-credal-set correspondence comes from a known class of hypotheses. We then define a notion of regret corresponding to the *lower prevision* defined by these credal sets. Equivalently, the setting can be regarded as a two-player *zero-sum game*, where, on each round, the agent chooses an arm and the adversary chooses the distribution over outcomes from a set of options associated with this arm. The regret is defined with respect to the value of game. For certain natural hypothesis classes, loosely analogous to stochastic linear bandits (which are a special case of the resulting setting), we propose an algorithm and prove a corresponding upper bound on regret.

Keywords: multi-armed bandits, game theory, regret bounds, upper confidence bound, imprecise probability

1. Introduction

The classical theory of multi-armed bandits (see e.g. Lattimore and Szepesvári (2020)) studies two main types of setting: stochastic and adversarial. In both cases, there is a sequence of rounds, and on each round an agent selects an "arm" and observes the resulting reward. In the stochastic setting, the reward is drawn from a distribution that depends only on the arm, independently from rewards on previous rounds. The obvious limitation is that real world problems often involve distributions that change over time and/or as a result of the agent's choices or external factors. In the adversarial setting, on each round, the reward is chosen for each arm in advance of the agent's decision (but not revealed to it), and the manner of choosing is arbitrary (it can even depend on the arms selected on previous arms). However, regret bounds in this setting require either bounding the number of arms, or assuming that reward depends in some regular manner on the arm (e.g. linearly). These assumptions are not always realistic.

We suggest making do with a weaker assumption, namely that the *set of possible outcome*[1] *distributions* has regular dependence on the arm. In other words, the set of responses available to

1. We say "outcome" and not "reward" because it turns out that, in this new setting, it is important to have auxiliary feedback in addition to the reward itself. See discussion in subsection 2.2.

the adversary changes with the arm in a regular manner, but, within the constraints of this set, the adversary's strategy can depend on the arm (more or less) arbitrarily.

The price we have to pay is remaining content with a weaker notion of regret, where we compare the expected reward to the maximin expected reward rather than to the best constant-arm response to the adversary's plays. This weaker notion of regret is natural for zero-sum games, where "strong" regret is often intractable (see e.g. Tian et al. (2021)). At the same time, this notion of regret can be regarded as the credal set analogue of regret in stochastic bandits.

Here is an example of a situation where our framework is applicable and classical approaches are not:

Example 1 *The algorithm needs to control the system of traffic lights in a city. On each day, the durations of the various states are set for each intersection. The resulting traffic is then measured, and its overall "quality" (e.g. a combination of travel time and number of accidents) is calculated. The traffic depends on hard to predict conditions that vary from day to day (e.g. events held around the city) and might be influenced by decisions on previous days (via changing driver habits). In particular it is not IID and cannot be modeled as a stochastic bandit. Moreover, a small change in the decision might trigger a large change in the outcome distribution (e.g. due to chaotic dynamics), implying it is not effective to model the problem as an adversarial bandit. On the other hand, we assume that finding a constant setting of traffic light durations that produces maximal expected quality under worst-case conditions is sufficient.*

We generalize stochastic bandits by allowing some "adversarial" degrees of freedom, but we treat these degrees of freedom differently than in classical adversarial bandits[2]. Our approach uses imprecise probability theory (see e.g. Augustin et al. (2014)). Imprecise probability theory provides ways to represent uncertainty more general than probability distributions. One such generalization is replacing the notion of a probability distribution by that of a *credal set*: a closed convex set of probability distributions. Instead of expected reward, we then consider the *minimal* expected reward over all distributions inside the set (the so-called "lower prevision": see Augustin et al. (2014)). Combining this formalism with that of stochastic linear bandits, produces a new framework that we call "imprecise linear bandits".

To apply our method to Example 1, we could, for instance, start with analytical modeling and/or simulations of traffic and arrive at some model which depends on unknown parameters, some of which don't substantially change over time whereas others can change unpredictably. We then treat the former as the "hypothesis" (unknown parameters on which the credal sets depend) and the latter as the adversarial choice of a distribution from within the credal set. For a more concrete demonstration, we give the following (extremely simplistic) example:

Example 2 *Suppose that our city has 3 roads: one from A (a residential neighborhood) to B (the beach), a second from A to C (the cinema) and a third from D to E. The DE road intersects the AB road at a single intersection and the AC road at a single intersection, so there are two intersections in total. Suppose also that nobody needs to travel e.g. from A to D, and therefore we can ignore the turns and each intersection has only two states. For instance, the ABDE intersection has a state in which the AB drivers have a green light and the DE drivers have a red light, and a state in which the AB drivers have a red light and the DE drivers have a green light.*

2. While classical stochastic bandits are a special case of our framework, classical adversarial bandits are *not* a special case of our framework.

The space of arms can then be parametrizes by the duration of each light at each intersection, i.e. it's

$$\mathcal{A} := [\tau_{\min}, \tau_{\max}]_{Bg} \times [\tau_{\min}, \tau_{\max}]_{Br} \times [\tau_{\min}, \tau_{\max}]_{Cg} \times [\tau_{\min}, \tau_{\max}]_{Cr}$$

Here $0 < \tau_{\min} < \tau_{\max}$ are the minimal and maximal durations a light is allowed to be, and the subscripts indicate which light: e.g. the subscript Bg means there is green light for the AB drivers whereas Cr means there is red light for the AC drivers.

The outcome is parameterized by the number of trips made along each road, which we model as a continuous number normalized to $[0, 1]$ since presumably these numbers are large. That is, the space of outcomes is

$$\mathcal{D} := [0, 1]_{AB} \times [0, 1]_{AC} \times [0, 1]_{DE}$$

The reward is minus the total waiting time of drivers on red lights, which we model as

$$r(x, y) := -\frac{1}{2}\left(\frac{x_{Br}^2 y_{AB} + x_{Bg}^2 y_{DE}}{x_{Bg} + x_{Br}} + \frac{x_{Cr}^2 y_{AC} + x_{Cg}^2 y_{DE}}{x_{Cg} + x_{Cr}}\right)$$

Here, x stands for the arm (i.e light durations) and y stands for the outcome (i.e. the traffic). This equation is derived by observing that each driver has a probability to encounter a red light equal to the fraction of time the light is red, and an average waiting time of half the duration of the red light (we assume there are no traffic jams or queues at intersections).

Now we need a model of driver behavior. Let's assume that the expected number of trips from D to E is an unknown constant $\theta_{DE} \in [0, 1]$. As to trips from A, we assume the residents require a constant amount of leisure in expectation, but can sometimes prefer the beach to the cinema or vice versa for hard to predict reasons (e.g. the weather, the movies currently available and the traffic light settings themselves which we assume to be known to them in advance of the trip). Hence the sum of trips from A to B and trips from A to C is in expectation equal to some unknown constant $\theta_A \in [0, 1]$. Thus, the hypothesis space is $\mathcal{H} := [0, 1]_{DE} \times [0, 1]_A$ and the credal sets are

$$\kappa_\theta := \{\zeta \in \Delta\mathcal{D} \mid \mathrm{E}_{y \sim \zeta}[y_{DE}] = \theta_{DE}, \mathrm{E}_{y \sim \zeta}[y_{AB} + y_{AC}] = \theta_A\}$$

The unpredicatable fluctuations of traffic between AB and AC means stochastic bandits are inapplicable. Moreover, the number of arms is infinite and the reward does not depend on the arm in a regular manner (e.g. it's possible that drivers abruptly change behavior when some light duration crosses a threshold: we make no assumptions about it), and hence adversarial bandits are also inapplicable. On the other hand, our main result (Theorem 2) produces a regret bound of $\tilde{O}(\sqrt{N})$ in this example, where N is the time horizon (and regret is defined w.r.t. the best reward achievable for worst-case admissible driver behavior).

For another example, consider the following:

Example 3 *A patient suffers from a condition with three primary symptoms: call them A, B and C. Every month, the doctor prescribes the patient treatment, by specifying the dosage of each of n different medicaments, between 0 to some maximal amount. The space of arms is hence $\mathcal{A} := [0, 1]^n$. In each month, the patient can either experience or not experience each of the 3 symptoms. The space of outcomes is hence $\mathcal{B} := 2^{\{A,B,C\}}$.*

Studies showed that the probability that a patient experiencing symptom A after receiving treatment x is given by $f^A(x)+\theta^A g^A(x)$, where $f^A, g^A : \mathcal{A} \to \mathbb{R}$ are fixed functions whereas $\theta^A \in [0,1]$ depends on the individual patient. Also, the probability of a patient experiencing symptom C conditional on each of the events (1) neither symptom A nor symptom B is present (2) symptom A is present but not symptom B (3) symptom B is present but not symptom A (4) both symptom A and symptom B are present, is given by $f_i^C(x) + \theta_i^C g_i^C(x)$, where $i \in \{1,2,3,4\}$ is the event, $f_i^C, g_i^C : \mathcal{A} \to \mathbb{R}$ are fixed functions and $\theta_i^C \in [0,1]$ depend on the individual patient. On the other hand, symptom B is difficult to predict and might depend both on the individual patient and on the patient's history in a complicated, poorly understood way.

The goal is maximizing the expected value of some function $r : \mathcal{A} \times \mathcal{B} \to \mathbb{R}$ that takes into account the severity of the various symptoms, and also the cost and long-term risks of the different medicaments, under worst-case assumptions about the unpredictable symptom B. Once again, stochastic bandits are inapplicable due to the nature of symptom B and adversarial bandits are inapplicable because the number of arms is infinite (and there is no good enough guarantee of smooth dependency). On the other hand, we still get a regret bound of $\tilde{O}(\sqrt{N})$ by Theorem 2.

Hopefully, future work will generalize this framework from bandits to reinforcement learning (i.e. allow a persistent observable state that is preserved over time), which would allow extending the possible range of applications much further.

Our framework can also be regarded as a repeated two-player zero-sum game, in which the arms are the agent's stage strategies, whereas a stage strategy of "nature" (the adversary) is an assignment to each arm of a distribution from that arm's respective credal set. Importantly, the agent doesn't know which credal set is associated with each arm (the same way as in stochastic bandits, it doesn't know which distribution is associated with each arm). The regret of an algorithm is then defined by comparing the minimal expected reward this algorithm guarantees with the minimal expected reward that could be guaranteed if the credal sets were known in advance.

Naturally, proving useful upper bounds on regret requires either bounding the number of arms, or assuming some structure on hypotheses (arm-to-credal-set assignments). We choose the latter path, since the former can be addressed using adversarial bandits, and the latter leads to a richer theory. Notice that, as opposed to adversarial bandits, the structure doesn't imply the reward depends on the arm in some regular way (e.g. linear, or even just Lipschitz), since the nature's policy can introduce irregularaties.

Specifically, we consider hypotheses classes with "linear" structure that generalizes the classical theory of linear bandits. In particular, this captures Example 2.

Related Work

Our framework can be naturally viewed as a generalization of stochastic linear bandits. For the latter, a regret bound optimal up to logarithmic factors was established in Dani et al. (2008).

It is also possible to view our framework as a zero-sum two-player game. Zero-sum two-player games with bandit feedback were studied in O'Donoghue et al. (2021). Indeed, their setting can be represented as a special case of our setting, including the notion of regret they chose to analyze. In O'Donoghue et al. (2021), each player has a finite set of arms (pure strategies in the stage game) and the regret bounds they show for their proposed algorithms are not better than the corresponding regret bound by Exp3 (however, they show that empirically Exp3 performs worse). On the other hand, our setting is much more general (in particular it allows for an infinite set of arms) and our regret bound is not obtained by any previously proposed algorithm (as far as we know).

Another work that studies learning zero-sum two-player games is Tian et al. (2021). Their notion of regret is also closely related to the present work. In one sense, the setting of Tian et al. (2021) is significantly more general than our setting, because they study (episodic) *stochastic* games: i.e. the game has a state, so it's more closely analogous to reinforcement learning than to multi-armed bandits. On the other hand, their regret bound scales with $|\mathcal{A}|^{\frac{1}{3}}$, where $\mathcal{A}$ is the set of actions (in particular, it has to be finite) and with $N^{\frac{2}{3}}$, where N is the number of episodes, in contrast to our bound (Theorem 2) which allows an infinite set of arms and scales with $N^{\frac{1}{2}}$ (where N is the time horizon). Therefore, these results are incomparable.

In Chen et al. (2022), a regret bound is proved for an algorithm that learns zero-sum two-player stochastic games which is roughly of the form $\tilde{O}(dN^{\frac{1}{2}})$, where d is the dimension of a certain linear feature space. However, they assume that *both* players are controlled by the algorithm (which is then required to converge to the Nash equilibrium of the game), whereas in our framework only one player is controlled by the algorithm, and the other is truly adversarial.

We also see a philosophical analogy between our approach and the notion of "semirandom models" in complexity theory (see Feige (2020)). Semirandom models interpolate between worst-case and average-complexity, by assuming the instances of a computational problem are drawn from a distribution which can be controlled in some limited, prescribed way by an adversary. Similarly, in our framework, the *outcomes* are sampled from a distribution which is controlled in some limited, prescribed way by an adversary. Semirandom models are concerned with computational complexity[3], while we are concerned with sample complexity (although the computational complexity of learning in this setting should also be studied.) In both cases the "adversary" is primarily viewed as a metaphor for properties of the real-world source of the data that are difficult or impossible to model explicitly (rather than an actual agent.)

The structure of the paper is as follows. Section 2 formally defines the setting and the notion of regret that we are going to analyze. Section 3 defines the IUCB algorithm. Section 4 proves upper bound on expected regret for IUCB. Section 5 studies the behavior of the parameters appearing in the regret bounds in some natural special cases. Section 6 summarizes the results and proposes directions for future work.

2. Setting

2.1 General Imprecise Bandits

We start by formally defining the framework. Fix $D_X \in \mathbb{N}$ and let $\mathcal{A}$ be a compact subset of $\mathbb{R}^{D_X}$ that represents arms[4]. Consider also a finite-dimensional vector space $\mathcal{Y}$, and fix a linear functional $\mu \in \mathcal{Y}^\star \setminus 0$ and a compact convex set $\mathcal{D} \subset \mu^{-1}(1) \subset \mathcal{Y}$ which represents outcomes[5]. (We use the notation * to denote the dual vector space.) $\mathcal{Y}$ and μ are purely technical devices for a convenient

3. Or with information-theoretic detection bounds: see e.g. Moitra et al. (2016) or Steinhardt (2017).
4. More generally, $\mathcal{A}$ can be an arbitrary compact metric space.
5. As opposed to classical multi-armed bandits, we will always work with some space of outcomes rather than just a reward scalar.

representation of $\mathcal{D}$. We assume the affine hull of $\mathcal{D}$ equals $\mu^{-1}(1)$. We also fix a continuous reward function $r : \mathcal{A} \times \mathcal{D} \to \mathbb{R}$.

Example 4 *An interesting special case is, when the outcomes that can be actually observed are some finite set $\mathcal{B}$ without any special structure. To represent this, we take $\mathcal{Y} := \mathbb{R}^{\mathcal{B}}$ and define μ and $\mathcal{D}$ by*

$$\mu(y) := \sum_{a \in \mathcal{B}} y_a$$

$$\mathcal{D} := \Delta\mathcal{B} = \left\{ y \in \mathbb{R}^{\mathcal{B}} \,\middle|\, \sum_{a \in \mathcal{B}} y_a = 1,\, \forall a \in \mathcal{B} : y_a \geq 0 \right\}$$

This is a valid representation since there is the canonical embedding $\iota : \mathcal{B} \to \Delta\mathcal{B}$ given by

$$\iota(a)_b := \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b \end{cases}$$

Given a reward function $r_0 : \mathcal{A} \times \mathcal{B} \to \mathbb{R}$, we can extend it to $r : \mathcal{A} \times \mathcal{D} \to \mathbb{R}$ as

$$r(x, y) := \sum_{a \in \mathcal{B}} y_a r_0(x, a)$$

A common way to represent beliefs in imprecise probability theory is *credal sets*. A credal set over $\mathcal{D}$ is a closed[6] convex subset of $\Delta\mathcal{D}$. We denote the space of credal sets by $\square\mathcal{D}$[7]. A *hypothesis* is a mapping $H : \mathcal{A} \to \square\mathcal{D}$. The meaning of this mapping is, whenever the agent pulls arm $x \in \mathcal{A}$, the outcome is drawn from some distribution in the set $H(x)$. The precise distribution is left unspecified: it can vary arbitrarily over time and as a function of the previous history of arms and outcomes, treated as adversarial. Given $x \in \mathcal{A}$ and H as above, the associated lower prevision (minimal expected reward) is

$$\mathrm{ME}_H\left[r|x\right] := \min_{\zeta \in H(x)} \mathrm{E}_{y \sim \zeta}\left[r(x, y)\right]$$

The optimal arm for H is then[8]

$$x^*_H := \underset{x \in \mathcal{A}}{\operatorname{argmax}}\, \mathrm{ME}_H\left[r|x\right]$$

If there are multiple optimal arms, the ambiguity can be resolved arbitrarily, except in particular cases we will point out.

An *agent policy* is a mapping[9] $\varphi : (\mathcal{A} \times \mathcal{D})^* \to \mathcal{A}$. That is, the agent chooses an action based on the history of previous actions and outcomes. A *nature policy* compatible with H is a mapping[10]

6. "Closed" in the sense of the weak topology on probability measures (see e.g. Klenke (2020), section 13.2).
7. Equipped with the Vietoris topology (see e.g. Beer (1993), section 2.2).
8. Notice that this is *not* the best arm in "hindsight" like in classical adversarial bandit theory. The notion of "hindsight" is not even well-defined because the adversary doesn't choose *counterfactual* outcomes, only actual. Rather, it is the "maximin" arm, the arm associated with the highest possible expected reward *guarantee* for the hypothesis H.
9. Technically, it has to be a Borel measurable mapping.
10. More precisely, a Markov kernel.

$\nu : (\mathcal{A} \times \mathcal{D})^* \times \mathcal{A} \to \Delta\mathcal{D}$ s.t. for any $(h, x) \in (\mathcal{A} \times \mathcal{D})^* \times \mathcal{A}$, $\nu(h, x) \in H(x)$. We denote by $\mathfrak{P}_H$ the set of all nature policies compatible with H. An agent policy φ together with a nature policy ν define a distribution $\varphi\nu \in \Delta((\mathcal{A} \times \mathcal{D})^\omega)$ in the natural way. That is, sampling $xy \sim \varphi\nu$ is performed recursively according to

$$\begin{cases} x_n & = \varphi\left(x_0 y_0 \ldots x_{n-1} y_{n-1}\right) \\ y_n & \sim \nu\left(x_0 y_0 \ldots x_{n-1} y_{n-1} x_n\right) \end{cases}$$

Given a policy φ, a hypothesis H and a time horizon $N \in \mathbb{N}$, the *expected regret* is

$$\mathrm{ERg}_H\left(\varphi; N\right) := N \cdot \mathrm{ME}_H\left[r | x_H^*\right] - \min_{\nu \in \mathfrak{P}_K} \sum_{n=0}^{N-1} \mathrm{E}_{xy \sim \varphi\nu}\left[r\left(x_n, y_n\right)\right]$$

Given also some $\delta \in (0, 1)$, the *δ-regret* is

$$\mathrm{Rg}_H\left(\varphi; N, \delta\right) := \min\left\{\rho \geq 0 \,\middle|\, \max_{\nu \in \mathfrak{P}_K} \Pr_{xy \sim \varphi\nu}\left[N \cdot \mathrm{ME}_H\left[r | x_H^*\right] - \sum_{n=0}^{N-1} r\left(x_n, y_n\right) \geq \rho\right] \leq \delta\right\}$$

That is, the δ-regret is the tightest upper bound on regret which holds with probability at least $1 - \delta$, for any nature policy.

Our goal is finding algorithms for which upper bounds on expected regret and δ-regret can be guaranteed for any hypothesis out of some class $\mathcal{HC}$. Notice that this is consistent with extending the outcome space, like in Example 4, because that strictly increases the set of possible nature policies, and hence any regret bound for the extended version also applies to the original version.

2.2 Linear Imprecise Bandits

Now, we spell out the exact assumptions we will use in the rest of the paper. We will make 4 such assumptions.

Assumption 1 *For any $H \in \mathcal{HC}$ and $x \in \mathcal{A}$, there is some closed convex subset $K(x)^+$ of $\mathcal{D}$ s.t.*

$$H(x) = \left\{\zeta \in \Delta\mathcal{D} \,\middle|\, \mathrm{E}_{y \sim \zeta}\left[y\right] \in K(x)^+\right\}$$

That is, we require that our credal set is defined entirely in terms of the *expected value* of the distribution.

Assumption 1 is a natural starting point for generalizing the classical theory of stochastic bandits. In stochastic bandits, the central assumption is that the distribution of reward for every arm is fixed over time. However, if we only assumed that the *expected values* of those distributions are fixed, we could derive only mildly weaker regret bounds: instead of a IID process we get a martingale, and instead of the Hoeffding inequality we can use the Azuma-Hoeffding inequality. This is essentially the special case of our setting where $K(x)^+$ consists of a single point.

In addition, Assumption 1 is less restrictive than it superficially seems, because additional moments of the distribution can be represented by embedding the outcome space into some higher-dimensional space. For example:

Example 5 *Consider some $n \geq 1$. Let $\mathcal{Y} := \mathbb{R}^{n+1}$, $\mu(y) := y_0$ and $\mathcal{D}$ be the convex hull of the moment curve $\mathfrak{M}$:*

$$\mathfrak{M} := \left\{ \begin{bmatrix} 1 \\ t \\ t^2 \\ \vdots \\ t^n \end{bmatrix} \middle| t \in [-1, +1] \right\}$$

The only actual feedback is a reward scalar $r \in [-1, +1]$, but we represent it as the point $y(r) \in \mathcal{D}$ given by $y_k(r) := r^k$. Then, Assumption 1 admits any credal set defined in terms of the first n moments *of the reward distribution.*

Assumption 2 *For any $H \in \mathcal{HC}$ and $x \in \mathcal{A}$, $K(x)^+$ is closed under affine (and not just convex) linear combinations. Equivalently, there exists $K(x)$ a linear subspace of $\mathcal{Y}$ s.t.*

$$K(x)^+ = K(x) \cap \mathcal{D}$$

The only motivation for this is that it makes deriving an upper regret bound much easier.

Superificially, Assumption 2 rules out the case in which the reward scalar is the only feedback, except for the special case of ordinary stochastic bandits. That's because if $\mathcal{D}$ is 1-dimensional then $K(x)^+$ has to be either a single point or all of $\mathcal{D}$. Indeed, simple examples of imprecise bandits in which each hypothesis is an *inequality* on the expected reward turn out to often have poor regret bounds[11]. However, our framework admits non-trivial examples with such feedback by embedding the reward into a higher-dimensional space.

Assumption 3 *Let $\mathcal{Y}$ be equipped with the norm whose unit ball is the absolute convex hull of $\mathcal{D}$. Then, r is convex and 1-Lipschitz (w.r.t. the norm on $\mathcal{Y}$) in the second (outcome) argument*[12].

Note that Assumptions 1 and 3 imply

$$\mathrm{ME}_H\left[r|x\right] = \min_{y \in K(x)^+} r(x, y)$$

In general, Lipschitz is a fairly mild condition, but convexity is not. However, given a reward function $r : \mathcal{A} \times \mathcal{D} \to \mathbb{R}$ which doesn't satisfy the convexity condition, we can construct the convex reward function

$$\tilde{r}(x, y) := \min_{\zeta \in \Delta\mathcal{D} : \mathrm{E}_{y' \sim \zeta}[y'] = y} \mathrm{E}_{y' \sim \zeta}\left[r(x, y')\right]$$

It is easy to see that $\tilde{r} \leq r$, and that for any $x \in \mathcal{A}$ and hypothesis H that satisfies Assumption 1,

$$\mathrm{ME}_H\left[\tilde{r}|x\right] = \mathrm{ME}_H\left[r|x\right]$$

11. We will not spell out the argument, but that is an obstacle that the author encountered on early attempts to find regret bounds for imprecise bandits.
12. Notice that r doesn't have to be Lipschitz in the arm argument, and the outcome can vary discontinuously with the arm if that's the adversary's strategy.

Hence, the expected regret for r is upper bounded by the expected regret for $\tilde{r}$, and any upper bound on the latter carries over to the former. The caveat is that $\tilde{r}$ might fail to be Lipschitz even if r is[13]. Even if $\tilde{r}$ is Lipschitz, its Lipschitz constant might be larger than the Lipschitz constant of r. This means that, even if r is 1-Lipschtiz and $\tilde{r}$ is Lipschitz, applying the results of this work might require rescaling $\tilde{r}$ by a constant multiplicative factor, which would then appear as an extra penalty in the regret bound (through the parameter C: see section 4). This caveat is why we cannot state the exact same results without the convexity condition.

Assumptions 1 and 2 are entirely about the individual credal sets. However, we also need to make an assumption about the relationship between different hypotheses (it is necessary even in classical theory). A common type of assumption in learning theory is requiring the hypothesis class is "low-dimensional" in some sense. In some cases, it takes the form of requiring an embedding of this class into a low-dimensional linear space in an appropriate way (in particular, this is the case for linear bandit theory). This is exactly the type of assumption we will me make.

Let $\mathcal{Z}$ be a finite-dimensional vector space, which will serve to parameterize our hypothesis class $\mathcal{HC}$, and $\mathcal{H} \subseteq \mathcal{Z}$ a compact set which will be the representation of $\mathcal{HC}$ within $\mathcal{Z}$. Let $\mathcal{W}$ be another finite-dimensional vector space, which will serve to parametrize the linear conditions that a credal set satisfying Assumptions 1 and 2 imposes on the expected value of the outcome distribution. That is, the $K(x)$ from Assumption 2 will be the kernel of some linear operator from $\mathcal{Y}$ to $\mathcal{W}$. Finally, let $F : \mathcal{A} \times \mathcal{Z} \times \mathcal{Y} \to \mathcal{W}$ be a mapping continuous in the first argument and bilinear in the second and third arguments, which serves to enable the interpretation of vectors in $\mathcal{Z}$ as hypotheses (see below). For any $x \in \mathcal{A}$ and $z \in \mathcal{Z}$, denote $F_{xz} : \mathcal{Y} \to \mathcal{W}$ the linear operator defined by $F_{xz}y := F(x, z, y)$. For any $x \in \mathcal{A}$ and $\theta \in \mathcal{H}$, denote $K_\theta(x) := \ker F_{x\theta}$.

For any $\mathcal{U} \subseteq \mathcal{Y}$ a linear subspace, denote $\mathcal{U}^+ := \mathcal{U} \cap \mathcal{D}$ and define $\kappa(\mathcal{U}) \in \square\mathcal{D}$ by

$$\kappa(\mathcal{U}) = \left\{ \zeta \in \Delta\mathcal{D} \,\middle|\, \mathrm{E}_{y\sim\zeta}[y] \in \mathcal{U}^+ \right\}$$

Assumption 4 *For any $x \in \mathcal{A}$ and $\theta \in \mathcal{H}$, $K_\theta(x) \cap \mathcal{D} \neq \varnothing$ and $F_{x\theta}$ is onto. Moreover,*

$$\mathcal{HC} = \{H : \mathcal{A} \to \square\mathcal{D} \mid \exists\theta \in \mathcal{H}\forall x \in \mathcal{A} : H(x) = \kappa(K_\theta(x))\}$$

Since $F_{x\theta}$ is onto, it follows that $\dim K_\theta(x) = \dim\mathcal{Y} - \dim\mathcal{W}$ and in particular it doesn't depend on x or θ. We will denote this number d.

Assumption 4 implies that any $\theta \in \mathcal{H}$ defines some $H \in \mathcal{HC}$, and conversly, any $H \in \mathcal{HC}$ is induced by some $\theta \in \mathcal{H}$. Given $\theta \in \mathcal{H}$ and its corresponding $H \in \mathcal{HC}$, we will use the shorthand notations $\mathrm{ME}_\theta := \mathrm{ME}_H$ and $\mathrm{ERg}_\theta := \mathrm{ERg}_H$.

Notice that the classical framework of stochastic linear bandits is a special case of linear imprecise bandits, which can be seen as follows.

Example 6 *Consider some $D \geq 1$, let $\mathcal{A} \subset \mathbb{R}^D$ and assume the linear span of $\mathcal{A}$ is all of $\mathbb{R}^D$. The only observable outcome is the reward which can be assumed to lie in $[-1, +1]$, so we take $\mathcal{Y} := \mathbb{R}^2$, $\mu = \begin{bmatrix} 1 & 0 \end{bmatrix}$, and*

$$\mathcal{D} := \left\{ \begin{bmatrix} 1 \\ t \end{bmatrix} \in \mathbb{R}^2 \,\middle|\, t \in [-1, +1] \right\}$$

13. However, if $\mathcal{D}$ is a polytope and r is 1-Lipschitz, then $\tilde{r}$ is c-Lipschitz, for some c that depends only on $\mathcal{D}$. The proof of this fact was suggested to the author by Felix Harder, and is not included in the present work.

Let $\mathcal{Z} := \mathbb{R}^D$ and $\mathcal{H}$ given by

$$\mathcal{H} := \left\{\theta \in \mathbb{R}^D \,\middle|\, \forall x \in \mathcal{A} : x^{\mathrm{t}}\theta \in [-1, +1]\right\}$$

The reward function is

$$r\left(x, \begin{bmatrix} 1 \\ t \end{bmatrix}\right) := t$$

Finally, let $\mathcal{W} := \mathbb{R}$ and set $F : \mathcal{A} \times \mathbb{R}^D \times \mathbb{R}^2 \to \mathbb{R}$ to be

$$F\left(x, z, \begin{bmatrix} y_0 \\ y_1 \end{bmatrix}\right) := (x^{\mathrm{t}} z)y_0 - y_1$$

It is easy to see we got the stochastic linear bandit with arm set $\mathcal{A}$.

In the rest of the paper, we use Assumptions 1-4 implicitly.

3. IUCB Algorithm

3.1 Norms

We propose an algorithm of UCB type which we call "imprecise UCB" (IUCB). The algorithm depends on a parameter $\eta \in \mathbb{R}_+$, the optimal choice of which will be discussed in section 4.

In order to describe the algorithm, we will need norms on the spaces $\mathcal{W}$ and a certain extension of $\mathcal{Z}$. We remind that the norm on $\mathcal{Y}$ is the unique norm s.t. the absolute convex hull of $\mathcal{D}$ is the unit ball. The norm on $\mathcal{W}$ is defined by

$$\|w\| := \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \min_{y \in \mathcal{Y} : F_{x\theta} y = w} \|y\|$$

Define the subspace $\mathcal{N} \subseteq \mathcal{Z} \oplus \mathcal{W}$ by

$$\mathcal{N} := \left\{ \begin{bmatrix} z \\ w \end{bmatrix} \in \mathcal{Z} \oplus \mathcal{W} \,\middle|\, \forall x \in \mathcal{A}, y \in \mathcal{Y} : F(x, z, y) + \mu(y)w = 0 \right\}$$

Denote $\bar{\mathcal{Z}} := (\mathcal{Z} \oplus \mathcal{W})/\mathcal{N}$. Define $\bar{F} : \mathcal{A} \times \bar{\mathcal{Z}} \times \mathcal{Y} \to \mathcal{W}$ by

$$\bar{F}\left(x, \begin{bmatrix} z \\ w \end{bmatrix} + \mathcal{N}, y\right) := F(x, z, y) + \mu(y)w$$

As before, $\bar{F}_{xz}$ is a linear operator from $\mathcal{Y}$ to $\mathcal{W}$.

Finally, the norm on $\bar{\mathcal{Z}}$ is defined by

$$\|z\| := \max_{x \in \mathcal{A}} \left\|\bar{F}_{xz}\right\|$$

Here, the norm on the right hand side the operator norm.

Notice that $\mathcal{W} \cap \mathcal{N} = \{0\}$, and therefore there is a natural embedding of $\mathcal{W}$ into $\bar{\mathcal{Z}}$. Moreover, the norm on $\mathcal{W}$ is compatible with the norm on $\bar{\mathcal{Z}}$ and this embedding, because $\|w \otimes \mu\| = \|w\| \cdot \|\mu\| = \|w\|$. We will also assume, without loss of generality, that $\mathcal{Z} \cap \mathcal{N} = \{0\}$, because otherwise we can just replace $\mathcal{Z}$ with $\mathcal{Z}/(\mathcal{Z}\cap\mathcal{N})$ and $\mathcal{H}$ with its image in $\mathcal{Z}/(\mathcal{Z}\cap\mathcal{N})$, without changing anything. Thereby, both $\mathcal{Z}$ and $\mathcal{W}$ will be viewed as *subspaces* of $\bar{\mathcal{Z}}$.

Extending $\mathcal{Z}$ to $\bar{\mathcal{Z}}$ is merely a device for defining the confidence set. Because, the constraint on the true hypothesis $\theta^* \in \mathcal{H}$ that can be inferred from observing a certain average outcome $\bar{y}$ after selecting a certain arm x multiple times, is most easily formulated as distance from a subspace in $\bar{\mathcal{Z}}$ (see the definition of $\mathcal{V}(x,\bar{y})$ below). And, $\mathcal{N}$ is just the "null" subspace that we quotient out because adding a vector in $\mathcal{N}$ has no effect on $\bar{F}$.

3.2 Algorithm

IUCB (Algorithm 1) works by maintaining a confidence set $\mathcal{C} \subseteq \mathcal{H}$ and applying the principle of optimism in the face of uncertainty. Initially, we set $\mathcal{C}_0 := \mathcal{H}$. On cycle $k \in \mathbb{N}$ of the algorithm, we choose the optimistic hypothesis θ_k:

$$\theta_k := \underset{\theta\in\mathcal{C}_k}{\operatorname{argmax}} \max_{x\in\mathcal{A}} \mathrm{ME}_\theta\left[r|x\right]$$

For any $x \in \mathcal{A}$ and $y \in \mathcal{Y}$, we will denote by $\bar{F}_x^y : \bar{\mathcal{Z}} \to \mathcal{W}$ the linear operator given by $\bar{F}_x^y z := \bar{F}(x,z,y)$. We also denote $\mathcal{V}(x,y) := \ker \bar{F}_x^y$. The algorithm selects arm $x^*_{\theta_k}$ *repeatedly* τ_k times, where τ_k is chosen to be minimal s.t.

$$\sqrt{\tau_k}\cdot \max_{\theta\in\mathcal{C}_k} \min_{z\in\mathcal{V}\left(x^*_{\theta_k},\bar{y}_k\right)} \|\theta - z\| \geq 2\left(D_Z+1\right)\eta \tag{1}$$

Here, $D_Z := \dim \mathcal{Z}$, $T_k := \sum_{i=0}^{k-1}\tau_k$, $\bar{y}_k := \frac{1}{\tau_k}\sum_{n=T_k}^{T_{k+1}-1} y_n$, where y_n is the outcome of round n, and $\eta \in \mathbb{R}_+$ is a parameter of the algorithm. That is, T_k is the total number of rounds of the first k cycles, and $\bar{y}_k$ is the average outcome over cycle k.

In other words, we keep selecting arm $x^*_{\theta_k}$ until condition (1) is met.

The idea behind equation (1) is, we select arm $x^*_{\theta_k}$ sufficiently many times so that the confidence set can be substantially narrowed down (by a factor of $\Omega(\eta D_Z)$) in directions traverse to $\mathcal{V}\left(x^*_{\theta_k},\bar{y}_k\right)$. We will use this to argue that such a narrowing down can only occure a small number of times ($\tilde{O}(D_Z^2)$) before the confidence set becomes thin (of size $O(D_Z^2 N^{-\frac{1}{2}})$) in those directions, and as a result the accrued expected regret becomes small.

After τ_k rounds, cycle k is complete and the confidence set is updated according to

$$\mathcal{C}_{k+1} := \left\{\theta\in\mathcal{C}_k \,\middle|\, \min_{z\in\mathcal{V}\left(x^*_{\theta_k},\bar{y}_k\right)} \|\theta - z\| \leq \frac{\eta}{\sqrt{\tau_k}}\right\}$$

Then, a new optimistic hypothesis θ_{k+1} is selected and we switch to selecting the arm $x^*_{\theta_{k+1}}$, et cetera.

4. Regret Bound

In addition to the dimensions of $\mathcal{Z}$, $\mathcal{Y}$ and $\mathcal{W}$, three parameters characterizing the hypothesis class are needed in order to formulate our regret bound.

The first is just

Algorithm 1 Imprecise UCB

Input $\eta \in \mathbb{R}_+$
$\mathcal{C} \leftarrow \mathcal{H}$
for k **from** 1 **to** ∞:
  $\theta^* \leftarrow \text{argmax}_{\theta \in \mathcal{C}} \max_{x \in \mathcal{A}} \text{ME}_\theta [r|x]$
  $\tau \leftarrow 0$
  $\Sigma y \leftarrow \mathbf{0} \in \mathcal{Y}$
  do
    select arm $x^*_{\theta^*}$ and observe outcome y
    $\tau \leftarrow \tau + 1$
    $\Sigma y \leftarrow \Sigma y + y$
    $\bar{y} \leftarrow \frac{\Sigma y}{\tau}$
  while $\sqrt{\tau} \cdot \max_{\theta \in \mathcal{C}} \min_{z \in \mathcal{V}\left(x^*_{\theta^*}, \bar{y}\right)} \|\theta - z\| < 2\left(D_Z + 1\right)\eta$
  $\mathcal{C} \leftarrow \left\{\theta \in \mathcal{C} \,\middle|\, \min_{z \in \mathcal{V}\left(x^*_{\theta^*}, \bar{y}\right)} \|\theta - z\| \leq \frac{\eta}{\sqrt{\tau}}\right\}$
end for

$$R := \max_{\theta \in \mathcal{H}} \|\theta\|$$

Notice that there is some redundancy in our description of a given bandit. First, there is redundancy in the choice of $\mathcal{H}$: given any $\theta \in \mathcal{Z}$ and $\chi \in \mathbb{R} \setminus 0$, $K_\theta = K_{\chi\theta}$. Hence, we get an equivalent bandit for any way of rescaling different hypotheses by different scalars. Second, there is redundancy in the choice of F: we can multiply it by any continuous function of x without affecting K_θ for any θ. While these redefinitions have no effect on the predictions different hypotheses make, and in particular no effect on regret, they *can* affect R. In order to get the tightest regret upper bound from our results, we need to choose the scaling with minimal R.

To define the second parameter we need the following:

Definition 1 *Let $\mathfrak{A}$ be a finite-dimensional affine space. Assume that the associated vector space $\vec{\mathfrak{A}}$ is equipped with a norm. Let $\mathcal{D} \subseteq \mathfrak{A}$ be a closed convex set and $\mathfrak{B} \subseteq \mathfrak{A}$ an affine subspace s.t. $\mathfrak{B} \cap \mathcal{D} \neq \varnothing$ and $\mathfrak{B} \not\subseteq \mathcal{D}$. We define the* sine *of $\mathfrak{B}$ relative to $\mathcal{D}$ to be*

$$\sin(\mathfrak{B}, \mathcal{D}) := \inf_{p \in \mathfrak{B} \setminus \mathcal{D}} \frac{\min_{q \in \mathcal{D}} \|p - q\|}{\min_{q \in \mathfrak{B} \cap \mathcal{D}} \|p - q\|}$$

Returning to our setting, for any linear subspace $\mathcal{U} \subseteq \mathcal{Y}$, we denote $\mathcal{U}^\flat := \mathcal{U} \cap \mu^{-1}(1)$. Then, the second parameter is

$$S := \min_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \sin\left(K_\theta(x)^\flat, \mathcal{D}\right)$$

While these parameters might seem complicated at first glance, the author's thesis (Kosoy (2024)) demonstrates that they are necessary (i.e. there are examples with provable lower bounds that preclude a power-law upper bound which doesn't depend on either R or S), and analyzes their behavior in various special cases.

Finally, the third parameter is just the size of the range of the reward function:

$$C := \max r - \min r$$

Notice that if r doesn't depend on the first argument, $C \leq 2$ because r is 1-Lipschitz.

We can now state our main result. Let $D_W := \dim \mathcal{W}$. Denote $\varphi^{\eta}_{\mathrm{IUCB}}$ the policy implemented by the IUCB algorithm.

Theorem 2 *Let N be a positive integer and fix any $\lambda > 0$. Denote*

$$\gamma := \frac{1}{\ln\left(1 - \frac{1}{e^2}\right)^{-1}}$$

Then, for all $\theta \in \mathcal{H}$

$$\begin{aligned}
\mathrm{ERg}_\theta\left(\varphi^{\eta}_{\mathrm{IUCB}}; N\right) \leq\, & 8\eta\left(S^{-1}+1\right) D_Z(D_Z+1)\sqrt{\gamma \ln \frac{D_Z R}{\lambda} \cdot N} \\
& + C D_W N^2 (N+1) \exp\left(-\Omega(1) \cdot \frac{\eta^2}{R^2 D_W^{\frac{5}{3}}}\right) \\
& + \gamma C D_Z^2 \ln \frac{D_Z R}{\lambda} \\
& + \left(S^{-1}+1\right) D_Z \left(36 D_Z + 8\right) N\lambda
\end{aligned}$$

In particular, we can set

$$\eta := \Theta(1) \cdot R D_W^{\frac{5}{6}} \sqrt{\ln\left(C D_W N\right)}$$

$$\lambda := \frac{1}{\sqrt{N}}$$

And then,

$$\mathrm{ERg}_\theta\left(\varphi^{\eta}_{\mathrm{IUCB}}; N\right) = \tilde{O}\left(R S^{-1} D_Z^2 D_W^{\frac{5}{6}} \sqrt{N} + D_Z^2 C\right)$$

Here and elsewhere, we use the notation $\tilde{O}(f) := O(f \cdot \mathrm{poly}(\ln f))$.

At first glance, it may appear strange that the bound in Theorem 2 doesn't depend on the size or dimension of $\mathcal{A}$. However, notice that we only "care" about arms of the form x^*_θ for some $\theta \in \mathcal{H}$. Indeed, other arms don't affect the definition of regret and are never selected in IUCB. Hence, effectively $\mathcal{A}$ can be assumed to be at most D_Z-dimensional.

Example 7 *In the setting of Example 6, $S = 1$, $R \leq 2$ (by Proposition 6), $D_Z = D$ and $D_W = 1$. Hence, Theorem 2 yields a regret bound of $\tilde{O}(D^2\sqrt{N})$. (As opposed to to minimax regret $\tilde{O}(D\sqrt{N})$, see Dani et al. (2008).)*

It is also possible to give a bound on δ-regret. We have,

Theorem 3 *In the setting of Theorem 2, for any $\delta > 0$ there is $\tilde{\delta} > 0$ s.t.*

$$\tilde{\delta} = \delta + D_W N(N+1)\exp\left(-\Omega(1)\cdot\frac{\eta^2}{R^2 D_W^{\frac{5}{3}}}\right)$$

and,

$$\begin{aligned}\mathrm{Rg}_\theta(\varphi_{\mathrm{IUCB}}; N, \tilde{\delta}) \leq & 8\eta\left(S^{-1}+1\right) D_Z(D_Z+1)\sqrt{\gamma\ln\frac{D_Z R}{\lambda}\cdot N} \\ & + C\sqrt{2N\ln\frac{1}{\delta}} \\ & + C D_W N^2(N+1)\exp\left(-\Omega(1)\cdot\frac{\eta^2}{R^2 D_W^{\frac{5}{3}}}\right) \\ & + \gamma C D_Z^2 \ln\frac{D_Z R}{\lambda} \\ & + \left(S^{-1}+1\right) D_Z\left(36 D_Z + 8\right) N\lambda\end{aligned}$$

In particular, we can set

$$\eta := \Theta(1)\cdot R D_W^{\frac{5}{6}}\sqrt{\ln\frac{C D_W N}{\delta}}$$

$$\lambda := \frac{1}{\sqrt{N}}$$

And then,

$$\mathrm{Rg}_\theta(\varphi_{\mathrm{IUCB}}; N, \delta) = \tilde{O}\left(\left(R S^{-1} D_Z^2 D_W^{\frac{5}{6}} + C\right)\sqrt{N\ln\frac{1}{\delta}} + D_Z^2 C\right)$$

The proofs of these theorems are in Appendix A.

5. Special Cases

5.1 Simple Bounds on R

Let $\mathcal{W} := \mathbb{R}$. In this case, $K_\theta(x)^\flat$ is a hyperplane inside $\mu^{-1}(1)$ for any $\theta \in \mathcal{H}$ and $x \in \mathcal{A}$. For any $x \in \mathcal{A}$, we can define the linear operator $A_x : \mathcal{Z} \to \mathcal{Y}^\star$ given by

$$\left(A_x z\right)(y) := F(x, z, y)$$

Assume that for all $x \in \mathcal{A}$, A_x is invertible (in particular, $\dim\mathcal{Y} = D_Z$). Let $\|\cdot\|_0$ be any norm on $\mathcal{Z}$ (which can be entirely different from the norm $\|\cdot\|$ which we defined in section 3). Together with the usual norm on $\mathcal{Y}$, this allows defining the operator norms $\|A_x\|_0$ and $\|A_x^{-1}\|_0$. We have,

Proposition 4 *In the setting above,*

$$R(\mathcal{H}, F) \leq \frac{\max_{\theta \in \mathcal{H}} \|\theta\|_0}{\min_{\theta \in \mathcal{H}} \|\theta\|_0} \left(\max_{x \in \mathcal{A}} \|A_x\|_0 \right) \left(\max_{a \in \mathcal{A}} \|A_x^{-1}\|_0 \right)$$

Here, we made the dependence of R on $\mathcal{H}$ and F explicit in the notation.

For any continuous function $\chi : \mathcal{H} \to \mathbb{R} \setminus 0$, we define

$$\mathcal{H}^\chi := \{\chi(\theta)\theta \mid \theta \in \mathcal{H}\}$$

Also, for any continuous function $\lambda : \mathcal{A} \to \mathbb{R} \setminus 0$, we define $F^\lambda : \mathcal{A} \times \mathcal{Z} \times \mathcal{Y} \to \mathcal{W}$ by

$$F^\lambda(x, z, y) := \lambda(x) F(x, z, y)$$

As we remarked in section 4, the bandit defined by $\mathcal{H}^\chi$ and F^λ is equivalent to that defined by $\mathcal{H}$ and F. However, it might have different R. We have

Proposition 5 *In the same setting, there exist $\chi : \mathcal{H} \to \mathbb{R} \setminus 0$ and $\lambda : \mathcal{A} \to \mathbb{R} \setminus 0$ s.t.*

$$R\left(\mathcal{H}^\chi, F^\lambda\right) \leq \max_{x \in \mathcal{A}} \left(\|A_x\|_0 \cdot \|A_x^{-1}\|_0\right)$$

This bound is an improvement on Proposition 4. Notice that the expression $\|A_x\|_0 \cdot \|A_x^{-1}\|_0$ is the *condition number* of A_x.

Now, we consider the case where $\mathcal{W} = \ker \mu$, and we have $\psi \in \mathcal{Z}^\star$ and $f : \mathcal{A} \times \mathcal{Z} \to \mathcal{Y}$ continuous and linear in the second argument s.t.

- For any $x \in \mathcal{A}$ and $z \in \mathcal{Z}$, $\mu(f(x, z)) = \psi(z)$.
- $\mathcal{H} \subseteq \psi^{-1}(1)$
- $f(\mathcal{A} \times \mathcal{H}) \subseteq \mathcal{D}$
- For any $x \in \mathcal{A}$, $z \in \mathcal{Z}$ and $y \in \mathcal{Y}$, $F(x, z, y) = \psi(z)y - \mu(y)f(x, z)$.

Then, for any $x \in \mathcal{A}$ and $\theta \in \mathcal{H}$, $K_\theta(x)^+ = \{f(x, \theta)\}$. Essentially, this datum is just an arm-dependent affine mapping from $\mathcal{H}$ to $\mathcal{D}$ that depends continuously on the arm. We have,

Proposition 6 *In the setting above,*

$$R(\mathcal{H}, F) \leq 2$$

See Appendix B for the proofs of the Propositions in this subsection.

5.2 Simple Bounds on S

For any $d \geq 1$, denote

$$\mathrm{Gr}_d^+(\mathcal{Y}) := \{\mathcal{U} \subseteq \mathcal{Y} \text{ a linear subspace} \mid \dim \mathcal{U} = d, \mathcal{U} \cap \mathcal{D} \neq \varnothing\}$$

$\mathrm{Gr}_d^+(\mathcal{Y})$ is a subset of the Grassmannian $\mathrm{Gr}_d(\mathcal{Y})$. Notice that $\mathrm{Gr}_1^+(\mathcal{Y})$ is canonically isomorphic to $\mathcal{D}$.

The following is a bound on S for the case where $\mathcal{D}$ is a simplex.

Proposition 7 *Let $\mathcal{B}$ be a finite set, $\mathcal{Y} := \mathbb{R}^{\mathcal{B}}$ and*

$$\mu(y) := \sum_{i \in \mathcal{B}} y_i$$

$$\mathcal{D} := \Delta\mathcal{B}$$

Consider any $d \geq 1$ and $\mathcal{U} \in \mathrm{Gr}_d^+(\mathcal{Y})$. Then,

$$\sin\left(\mathcal{U}^\flat, \mathcal{D}\right) \geq \max_{y \in \mathcal{U}^+} \min_{i \in \mathcal{B}} y_i$$

As we saw in Example 4, the simplex is a natural special case. We will use the perpsective of Example 4 throughout this section when discussing the simplex, even though the formal propositions are applicable more generally (i.e. they are still useful when non-vertex outcomes are allowed).

Proposition 7 allows us to lower bound the sine as long as the credal set has an outcome distribution with full support. More generally, we have the following

Proposition 8 *Let $\mathcal{B}$, $\mathcal{Y}$, μ, $\mathcal{D}$ and $\mathcal{U}$ be as in Proposition 7. Suppose that $\mathcal{E} \subseteq \mathcal{B}$ is s.t. $\mathcal{U} \subseteq \mathbb{R}^{\mathcal{E}}$. Then,*

$$\sin\left(\mathcal{U}^\flat, \mathcal{D}\right) \geq \max_{y \in \mathcal{U}^+} \min_{i \in \mathcal{E}} y_i$$

Here, $\mathbb{R}^{\mathcal{E}}$ is regarded as a subspace of $\mathbb{R}^{\mathcal{B}}$ by setting all the coordinates in $\mathcal{B} \setminus \mathcal{E}$ to 0.

This allows us to produce some lower bound for any given $\mathcal{U}$. Indeed, if $y_{1,2} \in \mathcal{U} \cap \Delta\mathcal{B}$, the support of y_1 is $\mathcal{E}_1$ and the support of y_2 is $\mathcal{E}_2$, then the support of $\frac{1}{2}(y_1 + y_2)$ is $\mathcal{E}_1 \cup \mathcal{E}_2$. Hence for any $\mathcal{U}$ there is some $\mathcal{E}$ s.t. all distributions in $\mathcal{U}$ are supported on it and some distribution in $\mathcal{U}$ has the entire $\mathcal{E}$ as its support.

Now, we consider the case where $\mathcal{D}$ is the unit ball.

Proposition 9 *Let $n \in \mathbb{N}$, $\mathcal{Y} = \mathbb{R}^{n+1}$ and*

$$\mu(y) := y_n$$

$$\mathcal{D} := \left\{ \begin{bmatrix} y \\ 1 \end{bmatrix} \in \mathbb{R}^{n+1} \,\middle|\, \|y\|_2 \leq 1 \right\}$$

Consider some $d \geq 1$ and $\mathcal{U} \in \mathrm{Gr}_d^+(\mathcal{Y})$. Then,

$$\sin\left(\mathcal{U}^\flat, \mathcal{D}\right) \geq \sqrt{1 - \min\left\{ \|y\|_2^2 \,\middle|\, \begin{bmatrix} y \\ 1 \end{bmatrix} \in \mathcal{U}^\flat \right\}}$$

This lower bound vanishes when $\mathcal{U}^\flat$ approaches a tangent to the ball: indeed, for a tangent, the sine is zero.

See Appendix C for the proofs of the Propositions in this subsection.

5.3 Probability Systems

Let's once again examine the case where $\mathcal{D}$ is a simplex and consider a credal set defined by fixing the probabilities of a family of events on $\mathcal{B}$.

Proposition 10 *Let $\mathcal{B}$ be a finite set, $\mathcal{Y} := \mathbb{R}^{\mathcal{B}}$ and $\mathcal{D} := \Delta\mathcal{B}$. Let $\mathcal{F}$ be a non-empty finite set and $f : \mathcal{B} \to 2^{\mathcal{F}}$ a surjection. Fix $p \in [0,1]^{\mathcal{F}}$ and define the linear subspace $\mathcal{U}$ of $\mathcal{Y}$ by*

$$\mathcal{U} := \left\{ y \in \mathcal{Y} \,\middle|\, \forall i \in \mathcal{F} : \sum_{a \in \mathcal{B}: i \in f(a)} y_a = p_i \sum_{a \in \mathcal{B}} y_a \right\}$$

Then,

$$\sin\left(\mathcal{U}^\flat, \mathcal{D}\right) \geq \frac{1}{|\mathcal{F}|}$$

In particular,

$$\sin\left(\mathcal{U}^\flat, \mathcal{D}\right) \geq \frac{1}{\log|\mathcal{B}|}$$

That is, each $i \in \mathcal{F}$ corresponds to an event $i \in f(a)$ whose probability we require to be p_i. The condition that f is surjective means that we require these events to be "logically independent" (i.e. it's possible for every one of them to happen or not, regardless of whether the others happened or not).

For $|\mathcal{F}| = 1$, we get $\sin\left(\mathcal{U}^\flat, \mathcal{D}\right) = 1$. The same holds if replace absolute probability by conditional probability:

Proposition 11 *Let $\mathcal{B}$ be a finite set, $\mathcal{Y} := \mathbb{R}^{\mathcal{B}}$ and $\mathcal{D} := \Delta\mathcal{B}$. Consider some $\mathcal{E}, \mathcal{E}' \subseteq \mathcal{B}$, fix $p \in [0,1]$ and define the linear subspace $\mathcal{U}$ of $\mathcal{Y}$ by*

$$\mathcal{U} := \left\{ y \in \mathcal{Y} \,\middle|\, \sum_{a \in \mathcal{E} \cap \mathcal{E}'} y_a = p \sum_{a \in \mathcal{E}} y_a \right\}$$

Then,

$$\sin\left(\mathcal{U}^\flat, \mathcal{D}\right) = 1$$

In the above, the subspace $\mathcal{U}$ can be interpreted as fixing the probability of event $\mathcal{E}'$ conditional on event $\mathcal{E}$ to the value p. Now we'll consider a *system* of conditional probabilities, but require a particular structure.

Proposition 12 *Fix an integer $n \geq 1$, and let $\mathcal{B} := \prod_{i<n} \mathcal{G}_i$ for some family of finite sets $\{\mathcal{G}_i\}_{i<n}$. Let $\mathcal{Y} := \mathbb{R}^{\mathcal{B}}$, $\mathcal{D} := \Delta\mathcal{B}$. For each $i < n$, denote*

$$\mathcal{G}_-^i := \prod_{j<i} \mathcal{G}_j$$

$$\mathcal{G}_+^i := \prod_{i<j<n} \mathcal{G}_j$$

Let $\mathcal{U}_i : \mathcal{G}_-^i \to \mathrm{Gr}^+ \left(\mathbb{R}^{\mathcal{G}_i}\right)$, where dropping the subscript d means we consider subspaces of all dimensions. Given $y \in \mathcal{Y}$, $i < n$ and $a \in \mathcal{G}_-^i$, let $\mathcal{E}_i(a) \subseteq \mathcal{G}_i$ be the minimal set s.t. $\mathcal{U}_i(a) \subseteq \mathbb{R}^{\mathcal{E}_i(a)}$ and define $y^a \in \mathbb{R}^{\mathcal{G}_i}$ by

$$y_b^a := \sum_{c \in \mathcal{G}_+^i} y_{abc}$$

Define the subspace $\mathcal{U}$ of $\mathcal{Y}$ by

$$\mathcal{U} := \{y \in \mathcal{Y} \mid \forall i < n, a \in \mathcal{G}_-^i : y^a \in \mathcal{U}_i(a) \text{ and } \forall b \in \mathcal{G}_i \setminus \mathcal{E}_i(a), c \in \mathcal{G}_+^i : y_{abc} = 0\}$$

Then,

$$\sin\left(\mathcal{U}^\flat, \Delta\mathcal{B}\right) \geq \min_{\substack{i<n \\ a \in \mathcal{G}_-^i}} \sin\left(\mathcal{U}_i(a)^\flat, \Delta\mathcal{G}_i\right)$$

That is, $\mathcal{U}_i(a)$ represents the set of admissible conditional probability distributions for the i-th component of the outcome given the condition a for the previous components. Notice that we define $\mathcal{U}$ s.t. any vector inside it vanishes outside the "support" of the $\mathcal{U}_i(a)$: this happens automatically for points in $\Delta\mathcal{B}$ but outside we could have cancellation between negative and positive components.

So, we can lower bound the sine of the "composite" subspace $\mathcal{U}$ by the minimum of the sines of the subspaces governing each conditional probability. Now, we'll show a similar bound for R.

In the setting of Proposition 12, suppose that we are given, for each $i < n$ and $a \in \mathcal{G}_-^i$, finite dimensional vector spaces $\mathcal{Z}_a$ and $\mathcal{W}_a$, a compact set $\mathcal{H}_a \subseteq \mathcal{Z}_a$ and $F_a : \mathcal{A} \times \mathcal{Z}_a \times \mathbb{R}^{\mathcal{G}_i} \to \mathcal{W}_a$, a mapping continuous in the first argument and bilinear in the second and third arguments. Assume that for any $x \in \mathcal{A}$ and $\theta \in \mathcal{H}_a$, $\ker F_{ax\theta} \cap \Delta\mathcal{G}_i \neq \varnothing$ and $F_{ax\theta}$ is onto. In other words, each a has its own bandit. Let $\mathcal{Z} := \bigoplus_{i,a} \mathcal{Z}_a$, $\mathcal{W} := \bigoplus_{i,a} \mathcal{W}_a$ and define F by

$$F(x, z, y)_a := F_a(x, z_a, y^a)$$

Define also $\mathcal{H} := \prod_{i,a} \mathcal{H}_a$. We have,

Proposition 13 *In the setting above*

$$R(\mathcal{H}, F) \leq 2n \max_{i,a} R(\mathcal{H}_a, F_a)$$

Example 8 *Consider Example 3. By Proposition 12, $S = 1$, since in this case $\mathcal{U}_i(a)$ is either 1-dimensional (for symptoms A and C) or the entire $\mathbb{R}^{\mathcal{G}_i}$ (for symptom B). By Proposition 13 and Proposition 6, $R \leq 12$ (since $n = 3$). Moreover, we have $D_Z = 10$ since there are 5 different θ parameters and each corresponds to some 2-dimensional $\mathcal{Z}_a$: this is the setting of Proposition 6 with $\mathcal{D}$ a finite closed interval describing symptom probability and $\mathcal{H}$ a finite closed interval describing one of the θ parameters. Finally, we have $D_W = 5$, since each θ parameter corresponds to some 1-dimensional $\mathcal{W}_a$ (in Proposition 6 for this case we have $\dim \mathcal{Y} = 2$ and hence $\dim \mathcal{W} = 1$).*

See Appendix D for the proofs of the Propositions in this subsection.

6. Summary

In this work, we proposed a new type of multi-armed bandit setting, importantly different from both stochastic and aversarial bandits, by replacing the probability distribution of a stochastic bandit by an imprecise belief (credal set). We studied a particular form of this setting, analogical to linear bandits. For this form, we proved an upper bound on regret of the usual form $\tilde{O}(\sqrt{N})$, where the coefficient includes the parameters S and R in addition to the more intuitive D_W, D_Z and C.

The author feels that we only scratched the surface of learing theory with imprecise probability. Here are a few possible directions for future research:

- We don't know whether the parameter D_W appearing in the upper bound is necessary. We can try to either prove a matching lower bound, or find an upper bound without D_W.
- For the special case of Example 6, Theorem 2 yields the regret bound $\tilde{O}(D^2\sqrt{N})$. However, we know from Dani et al. (2008) that a regret bound of $\tilde{O}(D\sqrt{N})$ is possible. It would be interesting to close this gap in a natural way. More generally, there is still significant room for tightening the gap between the lower bounds of Kosoy (2024) and the upper bound, in terms of the exponents of the parameters.
- Our analysis ignored computational complexity. It seems important to understand when IUCB or some other low regret algorithm can be implemented efficiently.
- The setting we study requires the hypothesis space to be naturally embedded in the vector space $\mathcal{Z}$ and scales with its dimension D_Z. It would be interesting to generalize this to regret bounds that scale with a *non-linear* dimension parameter, similarly to e.g. eluder dimension (see Russo and Van Roy (2013)) for stochastic bandits.
- Another possible direction is extending this setting from bandits to reinforcement learning. This seems very natural, since we can imagine an "MDP" transition kernel that takes values in credal sets instead of probability distributions. Indeed, the results of Tian et al. (2021) can already be interpeted in this way: but, they scale with the cardinalities of states and actions, whereas it would be better to have bounds in terms of dimension parameters.

Acknowledgments

This work was supported by the Machine Intelligence Research Institute in Berkeley, California, by Effective Ventures Foundation USA and by the Hebrew University of Jerusalem scholarship for outstanding students.

The author wishes to thank her advisor, Shai Shalev-Shwartz, for providing important advice during this work, and in particular emphasizing the need for examples. The author thanks Eran Malach for reading drafts and making useful suggestions for the improvement of Sections 1 and 2. The author thanks Felix Harder for solving a related technical question (see footnote 13). Finally, the author thanks Alexander Appel and her spouse Marcus Ogren for reading drafts, finding errors and making useful suggestions.

Appendix A. Proof of Regret Bound

The outline of the proof is as follows:

- For any $x \in \mathcal{A}$, $y \in \mathcal{D}$ and $\theta \in \mathcal{H}$, $d_Z(\theta, \mathcal{V}(x,y)) \leq R \cdot d_Y(y, K_\theta(x)^\flat)$, where d_Z stands for minimal distance w.r.t. the norm on $\bar{\mathcal{Z}}$.
- Let $\theta^* \in \mathcal{H}$ be the true hypothesis. For every cycle k and $\lambda > 0$, the probability that $d_Y(\bar{y}_k, K_{\theta^*}(x^*_{\theta_k})^\flat) \geq \lambda$ is upper bounded by $D_W \exp(-\Omega(1) \cdot D_W^{-\frac{5}{3}} \tau_k \lambda^2)$. This is proved with the help of the theorem from Giannopoulos (1995) and Azuma's inequality.
- As a consequence, θ^* always remains inside the confidence set $\mathcal{C}_k$ with high probability.
- The regret Reg_k accrued on cycle k is upper bounded by

$$\frac{1}{2}\tau_k d_Y\big(\bar{y}_k, K_{\theta_k}(x^*_{\theta_k})^+\big)$$

  Intuitively, if $\bar{y}_k$ is near $K_{\theta_k}(x^*_{\theta_k})^+$ then this is consistent with θ_k being close to θ^*, in which case we would suffer little regret.
- For any $y \in \mathcal{D}$, $\theta \in \mathcal{H}$ and $x \in \mathcal{A}$, $d_Y(y, K_\theta(x)^\flat) \leq O(1) \cdot d_Z(\theta, \mathcal{V}(x,y))$ and hence $d(y, K_\theta(x)^+) \leq O(1) \cdot S^{-1} d(\theta, \mathcal{V}(x,y))$. In particular, for $\theta \in \mathcal{C}_k$ we get the upper bound

$$O(1) \cdot S^{-1} \max_{\theta' \in \mathcal{C}_k} d_Z\big(\theta', \mathcal{V}(x,y)\big)$$

- Let $\bar{\mathcal{C}}_k$ be the convex hull of $\mathcal{C}_k$. Equation (1) guarantees that each cycle reduces the width of $\bar{\mathcal{C}}_k$ along some dimension by a factor of $2(D_Z + 1)$. This implies that the volume of $\bar{\mathcal{C}}_k$ is reduced by a factor of $\Omega(1)$.
- A D-dimensional convex body of volume at most $(D^{-1}\epsilon)^D \cdot v_D$ must have width at most 2ϵ along some dimension, where v_D is the volume of the unit ball. This is a consequence of John's ellipsoid theorem.
- It follows that after enough cycles, $\bar{\mathcal{C}}_k$ becomes "thin" along some dimension. This allows us to effectively go down to a proper subspace of $\mathcal{Z}$. The same reasoning applies in the lower dimension, leading to another reduction, et cetera, which can repeat no more than D_Z times. As a result, we can bound the number of cycles with substantial Reg_k. In these cycles, $\sum_k \mathrm{Reg}_k$ can be bounded using the inequalities above.

Now, let's dive into the details.

First, we want to bound distances in $\mathcal{Z}$ in terms of distances of compatible outcomes in $\mathcal{Y}$.

Lemma 14 *For any $x \in \mathcal{A}$, $y \in \mathcal{D}$ and $\theta \in \mathcal{H}$,*

$$d_Z(\theta, \mathcal{V}(x,y)) \leq R \cdot d_Y\Big(y, K_\theta(x)^\flat\Big)$$

Proof Denote $w := F(x, \theta, y)$. Then,

$$\begin{aligned}\bar{F}\left(x, \theta - w, y\right) &= F(x, \theta, y) - \mu(y) w \\ &= w - w \\ &= 0\end{aligned}$$

Therefore, $\theta - w \in \mathcal{V}\left(x, y\right)$ and $d_Z\left(\theta, \mathcal{V}\left(x, y\right)\right) \leq \|w\|$.

Let $y' \in K_\theta(x)^\flat$ be s.t. $\|y - y'\| = d_Y(y, K_\theta(x)^\flat)$. Then, $F_{x\theta} y' = 0$ and hence $w = F_{x\theta} y = F_{x\theta}(y - y')$. We get,

$$\begin{aligned}d_Z(\theta, \mathcal{V}\left(x, y\right)) &\leq \|w\| \\ &= \left\|F_{x\theta}\left(y - y'\right)\right\| \\ &\leq \|F_{x\theta}\| \cdot \left\|y - y'\right\| \\ &\leq \|\theta\| \cdot \left\|y - y'\right\| \\ &\leq R \cdot d_Y\Big(y, K_\theta(x)^\flat\Big)\end{aligned}$$

■

The following is a simple utility lemma:

Lemma 15 *For any $\mathcal{U} \in \mathrm{Gr}_d^+(\mathcal{Y})$ and $y \in \mathcal{D}$,*

$$d_Y(y, \mathcal{U}^\flat) \leq 4 d_Y(y, \mathcal{U})$$

Proof Denote $a := d_Y(y, \mathcal{U})$. Let $y^* \in \mathcal{U}$ be s.t. $\|y - y^*\| = a$. Notice that $\|y\| \leq 1$ and hence

$$\begin{aligned}\|y^*\| &\leq \|y\| + a \\ &\leq 1 + a\end{aligned}$$

Moreover, since $\|\mu\| = 1$ and $\mu(y) = 1$

$$\begin{aligned}\left|\mu\left(y^*\right) - 1\right| &= \left|\mu(y^*) - \mu(y)\right| \\ &\leq \|\mu\| \cdot \|y^* - y\| \\ &= a\end{aligned}$$

We consider 2 cases.

Case 1: $\mu(y^*) \geq \frac{1}{2}$. Then,

$$\left|1 - \frac{1}{\mu\left(y^*\right)}\right| \leq 1$$

Also,

$$\left|1-\frac{1}{\mu\left(y^{*}\right)}\right|=\frac{\left|\mu\left(y^{*}\right)-1\right|}{\left|\mu\left(y^{*}\right)\right|}$$
$$\leq\frac{a}{\left(\frac{1}{2}\right)}$$
$$=2a$$

Define $y^{\flat}\in\mathcal{U}^{\flat}$ by

$$y^{\flat}:=\frac{y^{*}}{\mu\left(y^{*}\right)}$$

We get

$$d_{Y}\left(y,\mathcal{U}^{\flat}\right)\leq\left\|y-y^{\flat}\right\|$$
$$\leq\left\|y-y^{*}\right\|+\left\|y^{*}-y^{\flat}\right\|$$
$$=a+\left|1-\frac{1}{\mu\left(y^{*}\right)}\right|\cdot\left\|y^{*}\right\|$$
$$\leq a+\min\left(1,2a\right)\cdot\left(1+a\right)$$
$$\leq a+3a$$
$$=4a$$

Case 2: $\mu(y^{*})<\frac{1}{2}$. Then, $a>\frac{1}{2}$ and

$$d_{Y}\left(y,\mathcal{U}^{\flat}\right)\leq\left\|y\right\|+d_{Y}\left(\mathbf{0},\mathcal{U}^{\flat}\right)$$
$$\leq 1+1$$
$$<4a$$

■

The next lemma uses a theorem from Giannopoulos (1995):

Theorem 16 (Giannopoulos) *Let $\mathcal{X}$ be a finite dimensional normed space. Denote $D:=\dim\mathcal{X}$. Then, the multiplicative Banach-Mazur distance from $\mathcal{X}$ to $\mathbb{R}^{D}$ equipped with the ℓ_1 norm*[14] *is at most $O(1)\cdot D^{\frac{5}{6}}$. That is, there exists a linear isomorphism $G:\mathbb{R}^{D}\xrightarrow{\sim}\mathcal{X}$ s.t. for all $x\in\mathbb{R}^{D}$*

$$\left\|x\right\|_{1}\leq\left\|Gx\right\|\leq O(1)\cdot D^{\frac{5}{6}}\left\|x\right\|_{1}$$

We now derive a concentration bound that shows the mean outcome over a sequence of rounds with constant arm x is close to $K_{\theta^{*}}(x)^{\flat}$ with high probability.

14. In Giannopoulos (1995), the theorem is stated for ℓ_{∞}, but it is the same by duality.

Lemma 17 *Suppose that arm $x \in \mathcal{A}$ is selected $\tau \in \mathbb{N}$ times in a row. Denote the average outcome by $\bar{y} \in \mathcal{D}$ and let $\delta > 0$. Then,*

$$\Pr\left[d_Y\Big(\bar{y}, K_{\theta^*}(x)^\flat\Big) \geq \delta\right] \leq 2D_W \exp\left(-\Omega(1) \cdot \frac{\tau\delta^2}{D_W^{\frac{5}{3}}}\right) \tag{2}$$

Proof Let $\mathcal{U} := K_{\theta^*}(x)$, $\mathcal{X} := (\mathcal{Y}/\mathcal{U})^\star$. $\mathcal{X}$ is a subspace of $\mathcal{Y}^\star$, and therefore has a natural norm. Observe that $\dim \mathcal{X} = D_W$, and apply Theorem 16 to get $G : \mathbb{R}^{D_W} \to \mathcal{X}$. Let T be the time our sequence of rounds begins, so that for each $n < \tau$, y_{T+n} is the outcome of the n-th round in the sequence. For each $i < D_W$, let $e_i \in \mathbb{R}^{D_W}$ be the i-th canonical basis vector and $\alpha_i := Ge_i$. Then, $\alpha_i(y_{T+n})$ is a random variable whose range is contained in an interval of length

$$\left|\max_{y \in \mathcal{D}} \alpha_i(y) - \min_{y \in \mathcal{D}} \alpha_i(y)\right| \leq 2\left\|\alpha_i\right\| \leq O(1) \cdot D_W^{\frac{5}{6}}$$

By the Azuma-Hoeffding inequality,

$$\Pr\left[\left|\alpha_i(\bar{y})\right| \geq \delta\right] \leq 2\exp\left(-\Omega(1) \cdot \frac{\tau\delta^2}{D_W^{\frac{5}{3}}}\right)$$

Applying a union bound, we get

$$\Pr\left[\max_i \left|\alpha_i(\bar{y})\right| \geq \delta\right] \leq 2D_W \exp\left(-\Omega(1) \cdot \frac{\tau\delta^2}{D_W^{\frac{5}{3}}}\right)$$

Observe that

$$\begin{aligned}
d_Y(\bar{y}, \mathcal{U}) &= \max_{\beta \in \mathcal{X} : \|\beta\| \leq 1} \beta(\bar{y}) \\
&= \max_{v \in \mathbb{R}^{D_W} : \|Gv\| \leq 1} Gv(\bar{y}) \\
&\leq \max_{v \in \mathbb{R}^{D_W} : \|v\|_1 \leq 1} Gv(\bar{y}) \\
&= \max_{v \in \mathbb{R}^{D_W} : \|v\|_1 \leq 1} \sum_{i < D_W} v_i \alpha_i(\bar{y}) \\
&= \max_{i < D_W} \left|\alpha_i(\bar{y})\right|
\end{aligned}$$

Hence,

$$\Pr\left[d_Y(\bar{y}, \mathcal{U}) \geq \delta\right] \leq 2D_W \exp\left(-\Omega(1) \cdot \frac{\tau\delta^2}{D_W^{\frac{5}{3}}}\right)$$

Since $\bar{y} \in \mathcal{D}$, Lemma 15 implies that $d_Y(\bar{y}, \mathcal{U}^\flat) \leq 4d_Y(\bar{y}, \mathcal{U})$. Redefining δ, we get (2). ■

From now on, we'll use the notation $\Pr[A; B]$ to mean "probability of the event A and B" and

$\mathrm{E}[f;B]$ to mean "the expected value of $f \cdot \mathbf{1}_B$, where $\mathbf{1}_B$ is defined to be 1 when event B happens and 0 when event B doesn't happen".

The next lemma shows that, given a concentration bound such as in Lemma 17, the true hypothesis remains in the confidence set with high probability.

Lemma 18 *Assume that a given bandit is s.t., in the setting of Lemma 17, it always holds that*

$$\Pr\left[d_Y\Big(\bar{y}, K_{\theta^*}(x)^\flat\Big) \geq \delta\right] \leq f\left(\tau\delta^2\right)$$

Here, $f : \mathbb{R} \to \mathbb{R}$ is some continuous function.

Assume agent policy $\varphi^\eta_{\mathrm{IUCB}}$ and any *nature policy. Then, for any $N \in \mathbb{N}$*

$$\Pr\left[\exists k : T_k < N \textit{ and } \theta^* \notin \mathcal{C}_k\right] \leq \frac{1}{2}N(N+1)f\left(\frac{\eta^2}{R^2}\right)$$

Proof For any $i, j \in \mathbb{N}$ s.t. $i < j$, denote $\bar{y}_{ij} := \frac{1}{j-i}\sum_{n=i}^{j-1} y_n$. Denote A_{ij} the event $\forall i \leq n < j : x_n = x_i$. By the assumption,

$$\Pr\left[d_Y\Big(\bar{y}_{ij}, K_{\theta^*}(x_i)^\flat\Big) \geq \delta; A_{ij}\right] \leq f\left((j-i)\delta^2\right)$$

Using Lemma 14, we get

$$\Pr\left[\frac{1}{R} \cdot d_Z(\theta^*, \mathcal{V}(x_i, \bar{y}_{ij})) \geq \delta; A_{ij}\right] \leq f\left((j-i)\delta^2\right)$$

Substituting $\delta := \frac{1}{R} \cdot \frac{\eta}{\sqrt{j-i}}$, this becomes

$$\Pr\left[d_Z(\theta^*, \mathcal{V}(x_i, \bar{y}_{ij})) \geq \frac{\eta}{\sqrt{j-i}}; A_{ij}\right] \leq f\left(\frac{\eta^2}{R^2}\right)$$

Denote B_{ij} the event inside the probability operator on the left hand side. Taking a union bound, we get

$$\Pr\left[\exists 0 \leq i < j \leq N : B_{ij}\right] \leq \frac{N(N+1)}{2} \cdot f\left(\frac{\eta^2}{R^2}\right)$$

From the definition of $\mathcal{C}_k$, it is clear that if $\forall 0 \leq i < j \leq N$: not-B_{ij}, then $\forall k$: if $T_k < N$ then $\theta^* \in \mathcal{C}_k$. ■

Putting Lemma 17 and Lemma 18 together immediately gives us:

Lemma 19 *Assume agent policy $\varphi^\eta_{\mathrm{IUCB}}$ and* any *nature policy. Then, for any $N \in \mathbb{N}$*

$$\Pr\left[\exists k : T_k < N \textit{ and } \theta^* \notin \mathcal{C}_k\right] \leq D_W N(N+1)\exp\left(-\Omega(1) \cdot \frac{\eta^2}{R^2 D_W^{\frac{5}{3}}}\right)$$

Now we show that the regret of a sequence of rounds with an optimistically selected arm can be bounded in terms of the minimal distance of the mean outcome from the possible expected outcomes of the optimistic hypothesis. We denote $r^* := \mathrm{ME}_{\theta^*}\left[r|x^*_{\theta^*}\right]$.

Lemma 20 *Suppose $\theta \in \mathcal{H}$ is s.t.*

$$\mathrm{ME}_{\theta}\left[r|x^*_{\theta}\right] \geq r^* \tag{3}$$

Assume that the action $x^*_\theta \in \mathcal{A}$ is selected $\tau \in \mathbb{N}$ times in a row, starting from round $T \in \mathbb{N}$. Denote the average outcome by $\bar{y}$. Then,

$$\tau r^* - \sum_{n<\tau} r\left(x^*_\theta, y_{T+n}\right) \leq \tau d_Y\big(\bar{y}, K_\theta\left(x^*_\theta\right)^+\big) \tag{4}$$

Proof By convexity of r,

$$\frac{1}{\tau}\sum_{n<\tau} r\left(x^*_\theta, y_{T+n}\right) \geq r\left(x^*_\theta, \bar{y}\right)$$

Since r is 1-Lipschitz, this implies

$$\begin{aligned}\frac{1}{\tau}\sum_{n<\tau} r\left(x^*_\theta, y_{T+n}\right) &\geq \min_{y \in K_\theta\left(x^*_\theta\right)^+} r\left(x^*_\theta, y\right) - d_Y\big(\bar{y}, K_\theta\left(x^*_\theta\right)^+\big) \\ &= \mathrm{ME}_\theta\left[r|x^*_\theta\right] - d_Y\big(\bar{y}, K_\theta\left(x^*_\theta\right)^+\big)\end{aligned}$$

Using (3) and rearranging, we get (4). ■

The following is a simple geometric observation through which "sines" make it into the regret bound.

Lemma 21 *Let $\mathfrak{A}$ be a finite-dimensional affine space. Assume that the associated vector space $\vec{\mathfrak{A}}$ is equipped with a norm. Let $\mathcal{D} \subseteq \mathfrak{A}$ be a closed convex set and $\mathfrak{B} \subseteq \mathfrak{A}$ an affine subspace. Assume $\sin\left(\mathfrak{B}, \mathcal{D}\right) > 0$ and consider any $y \in \mathcal{D}$. Then,*

$$d\left(y, \mathfrak{B} \cap \mathcal{D}\right) \leq \left(\frac{1}{\sin\left(\mathfrak{B}, \mathcal{D}\right)} + 1\right) d\left(y, \mathfrak{B}\right) \tag{5}$$

Proof Let $y' \in \mathfrak{B}$ be s.t. $\|y - y'\| = d\left(y, \mathfrak{B}\right)$. If $y' \in \mathcal{D}$ then $d(y, \mathfrak{B} \cap \mathcal{D}) = d(y, \mathfrak{B})$ and (5) is true. Now, suppose $y' \notin \mathcal{D}$. Then, by definition of $\sin\left(\mathfrak{B}, \mathcal{D}\right)$,

$$\frac{d\left(y', \mathcal{D}\right)}{d\left(y', \mathfrak{B} \cap \mathcal{D}\right)} \geq \sin\left(\mathfrak{B}, \mathcal{D}\right)$$

We get

$$
\begin{aligned}
d(y, \mathfrak{B} \cap \mathcal{D}) &\leq \left\|y - y'\right\| + d\left(y', \mathfrak{B} \cap \mathcal{D}\right) \\
&\leq \left\|y - y'\right\| + \frac{d(y', \mathcal{D})}{\sin(\mathfrak{B}, \mathcal{D})} \\
&\leq \left\|y - y'\right\| + \frac{\|y' - y\|}{\sin(\mathfrak{B}, \mathcal{D})} \\
&= \left(1 + \frac{1}{\sin(\mathfrak{B}, \mathcal{D})}\right) d(y, \mathfrak{B})
\end{aligned}
$$

■

Now we derive a sort of converse to Lemma 14.

Lemma 22 *For any $y \in \mathcal{D}$, $\theta \in \mathcal{H}$ and $x \in \mathcal{A}$*

$$
d_Y\Big(y, K_\theta(x)^\flat\Big) \leq 4 d_Z\left(\theta, \mathcal{V}(x, y)\right) \tag{6}
$$

Proof Let $z \in \mathcal{V}(x, y)$ be s.t. $\|\theta - z\| = d_Z(\theta, \mathcal{V}(x, y))$. Denote $w := F(x, \theta, y)$. Since $\bar{F}(x, z, y) = 0$, we have $w = \bar{F}(x, \theta - z, y)$. Therefore

$$
\begin{aligned}
\|w\| &= \left\|\bar{F}_{x,\theta - z} y\right\| \\
&\leq \left\|\bar{F}_{x,\theta - z}\right\| \cdot \|y\| \\
&\leq \|\theta - z\| \cdot 1 \\
&= d_Z(\theta, \mathcal{V}(x, y))
\end{aligned}
$$

By definition of the norm on $\mathcal{W}$, there exists $\delta y \in \mathcal{Y}$ s.t. $F_{x\theta} \delta y = w$ and $\|\delta y\| \leq \|w\|$. Denote $y' := y - \delta y$. We have

$$
\begin{aligned}
F_{x\theta} y' &= F_{x\theta} y - F_{x\theta} \delta y \\
&= w - w \\
&= 0
\end{aligned}
$$

Hence, $y' \in K_\theta(x)$ and therefore

$$
\begin{aligned}
d_Y(y, K_\theta(x)) &\leq \left\|y - y'\right\| \\
&\leq \|w\| \\
&\leq d_Z(\theta, \mathcal{V}(x, y))
\end{aligned}
$$

Applying Lemma 15, we get (6). ■

The next lemma shows that when a D-dimensional convex set is "sliced" in a manner which reduces its width along some dimension by a factor of $\Omega(D)$, its volume must become reduced by a factor of $\Omega(1)$. This is the key to exploiting the finite dimension of the hypothesis space in the regret bound. The idea of the proof is reducing the problem to the special case of a cone sliced parallel to its base, in which case the conclusion follows from symmetry.

Lemma 23 *Let $D \geq 1$, $\mathcal{X}$ a D-dimensional vector space, $\alpha \in \mathcal{X}^\star$ and $\mathcal{C} \subseteq \mathcal{X}$ a compact convex set s.t.*

$$\max_{z \in \mathcal{C}} |\alpha(z)| = 1$$

Define $\mathcal{C}' \subseteq \mathcal{X}$ by

$$\mathcal{C}' := \left\{ z \in \mathcal{C} \,\middle|\, |\alpha(z)| \leq \frac{1}{D+1} \right\}$$

Then,

$$vol(\mathcal{C}') \leq \left(1 - \frac{1}{e^2}\right) \cdot vol(\mathcal{C}) \tag{7}$$

Here, vol stands for volume. Notice that volume is only defined up to a scalar, but since it appears on both sides of the inequality, it doesn't matter.

Proof [Proof of Lemma 23]Assume w.l.o.g. that there is $v \in \mathcal{C}$ s.t. $\alpha(v) = +1$, and fix some v like that. Define $\mathcal{B}_+, \mathcal{B}_- \subseteq \mathcal{X}$ by

$$\mathcal{B}_+ := \left\{ z \in \mathcal{C} \,\middle|\, \alpha(z) = +\frac{1}{D+1} \right\}$$
$$\mathcal{B}_- := \left\{ \left(1 + \frac{2}{D}\right) z - \frac{2}{D} v \,\middle|\, z \in \mathcal{B}_+ \right\}$$

For any $z \in \mathcal{B}_-$, we have

$$\begin{aligned}
\alpha(z) &= \alpha\left(\left(1 + \frac{2}{D}\right) z' - \frac{2}{D} v\right) \\
&= \left(1 + \frac{2}{D}\right) \alpha\left(z'\right) - \frac{2}{D} \alpha(v) \\
&= \left(1 + \frac{2}{D}\right) \frac{1}{D+1} - \frac{2}{D} \\
&= \frac{1}{D+1} \left(1 + \frac{2}{D} - \frac{2\,(D+1)}{D}\right) \\
&= \frac{1}{D+1} \cdot \frac{D + 2 - 2\,(D+1)}{D} \\
&= \frac{1}{D+1} \cdot \frac{-D}{D} \\
&= -\frac{1}{D+1}
\end{aligned}$$

Here, z' is some point in $\mathcal{B}_+$.

Let $\mathcal{C}_\pm$ be the convex hulls of $\mathcal{B}_\pm$ and v. It is easy to see that $\mathcal{C}_+ \setminus \mathcal{B}_+ \subseteq \mathcal{C} \setminus \mathcal{C}'$ and $\mathcal{C}' \setminus \mathcal{B}_+ \subseteq \mathcal{C}_- \setminus \mathcal{C}_+$. Moreover, $\mathcal{C}_-$ can be obtained from $\mathcal{C}_+$ by a homothety with center v and ratio $1 + \frac{2}{D}$, and hence

$$
\begin{aligned}
vol\left(\mathcal{C}_{-}\right) &= \left(1+\frac{2}{D}\right)^{D} vol\left(\mathcal{C}_{+}\right) \\
&\leq e^{2} \cdot vol\left(\mathcal{C}_{+}\right)
\end{aligned}
$$

It follows that

$$
\begin{aligned}
vol(\mathcal{C}') &\leq vol(\mathcal{C}_{-}) - vol(\mathcal{C}_{+}) \\
&\leq e^{2} vol(\mathcal{C}_{+}) - vol(\mathcal{C}_{+}) \\
&= \left(e^{2}-1\right) vol(\mathcal{C}_{+}) \\
&\leq \left(e^{2}-1\right)\left(vol(\mathcal{C}) - vol(\mathcal{C}')\right)
\end{aligned}
$$

Rearranging, we get (7). ■

The following allows us to lower bound the volume of a convex body in terms of its *minimal width*. Combined with Lemma 23, it will allow us to ensure the confidence set effectively becomes lower dimensional over time.

Lemma 24 *Let $D \geq 1$ and $\mathcal{C} \subseteq \mathbb{R}^D$ a compact convex set. Define $w \in [0, \infty)$ by*

$$
w := \min_{\alpha \in \mathbb{R}^D : \|\alpha\|_2 = 1} \left(\max_{z \in \mathcal{C}} \alpha^t z - \min_{z \in \mathcal{C}} \alpha^t z \right)
$$

Then,

$$
vol(\mathcal{C}) \geq \left(\frac{w}{2D}\right)^{D} v_D
$$

Here, v_D is the volume of the unit ball in $\mathbb{R}^D$.

Proof [Proof of Lemma 24]By John's ellipsoid theorem[15], there is an ellipsoid[16] $\mathcal{E}$ in $\mathbb{R}^D$ with the following property. Denote by $\mathcal{E}'$ the ellipsoid obtained from $\mathcal{E}$ by a homothety with center at the center of $\mathcal{E}$ and ratio $\frac{1}{D}$. Then, $\mathcal{E}' \subseteq \mathcal{C} \subseteq \mathcal{E}$. Each axis of $\mathcal{E}$ must be $\geq w$, hence each axis of $\mathcal{E}'$ must be $\geq \frac{w}{D}$, and therefore

$$
\begin{aligned}
vol(\mathcal{C}) &\geq vol(\mathcal{E}') \\
&\geq \left(\frac{w}{2D}\right)^{D} v_D
\end{aligned}
$$

■

Lemmas 23 and 24 will be used to show that the confidence set effectively becomes lower dimensional, in the sense that we can bound its thickness along multiple dimensions. However, the "slicing" performed by IUCB still happens in the original D_Z-dimensional space rather than the effective subspace. In order to translate the slicing factor from the ambient space to the subspace, we will use the following.

15. See Giorgi and Kjeldsen (2013), p. 213, Theorem III
16. By an "ellipsoid" we mean the convex body, rather than the surface. That is, an ellipsoid is defined to be the image of a closed ball under an affine transformation.

Lemma 25 *Let $D \geq 1$, $E \geq 0$, $\lambda \in (0, \infty)$, $\mathcal{C} \subseteq \mathbb{R}^D \times [-\lambda, +\lambda]^E$ a compact set and $\alpha \in \mathbb{R}^{D+E}$ s.t. $\|\alpha\|_2 \leq 1$. Define $\rho_0 \in (0, \infty)$ by*

$$\rho_0 := \max_{z \in \mathcal{C}} \left|\alpha^{\mathrm{t}} z\right|$$

Define $\mathcal{C}' \subseteq \mathbb{R}^{D+E}$ by

$$\mathcal{C}' := \left\{ z \in \mathcal{C} \,\middle|\, \left|\alpha^{\mathrm{t}} z\right| \leq \frac{1}{2(D+1)} \cdot \rho_0 \right\}$$

Let $P : \mathbb{R}^{D+E} \to \mathbb{R}^D$ be the orthogonal projection, and define $\rho_1, \rho_1' \in (0, \infty)$ by

$$\begin{aligned}
\rho_1 &:= \max_{z \in \mathcal{C}} \left|\alpha^{\mathrm{t}} P z\right| \\
\rho_1' &:= \max_{z \in \mathcal{C}'} \left|\alpha^{\mathrm{t}} P z\right|
\end{aligned}$$

Assume $\rho_0 \geq (4D + 5)\sqrt{E} \cdot \lambda$. Then,

$$\rho_1' \leq \frac{1}{D+1} \cdot \rho_1$$

Proof For any $z \in \mathbb{R}^D \times [-\lambda, +\lambda]^E$, we have

$$\alpha^{\mathrm{t}} z - \sqrt{E} \cdot \lambda \leq \alpha^{\mathrm{t}} P z \leq \alpha^{\mathrm{t}} z + \sqrt{E} \cdot \lambda$$

It follows that

$$\begin{aligned}
\rho_1 &\geq \rho_0 - \sqrt{E} \cdot \lambda \\
&\geq (4D + 5)\sqrt{E} \cdot \lambda - \sqrt{E} \cdot \lambda \\
&= 4(D + 1)\sqrt{E} \cdot \lambda
\end{aligned}$$

Moreover,

$$\begin{aligned}
\rho_1' &\leq \frac{1}{2(D+1)} \cdot \rho_0 + \sqrt{E} \cdot \lambda \\
&\leq \frac{1}{2(D+1)} \cdot \left(\rho_1 + \sqrt{E} \cdot \lambda\right) + \sqrt{E} \cdot \lambda \\
&= \frac{1}{2(D+1)} \cdot \rho_1 + \left(1 + \frac{1}{2(D+1)}\right) \sqrt{E} \cdot \lambda \\
&\leq \frac{1}{2(D+1)} \cdot \rho_1 + \left(1 + \frac{1}{2(D+1)}\right) \frac{1}{4(D+1)} \cdot \rho_1 \\
&\leq \frac{1}{2(D+1)} \cdot \rho_1 + 2 \cdot \frac{1}{4(D+1)} \cdot \rho_1 \\
&= \frac{1}{D+1} \cdot \rho_1
\end{aligned}$$

■

We also need to bound the regret once the confidence set effectively becomes 0-dimensional, i.e. is contained within a small hypercube. For this we use the following.

Lemma 26 *Suppose $\theta, \theta^* \in \mathcal{H}$ are s.t.*

$$\mathrm{ME}_\theta \left[r | x^*_\theta \right] \geq r^* \tag{8}$$

Let y be a $\mathcal{D}$-valued random variable whose distribution lies in $\kappa(K_{\theta^*}(x^*_\theta))$. Then,

$$r^* - \mathrm{E}\left[r\left(x^*_\theta, y \right) \right] \leq 4 \left(S^{-1} + 1 \right) \| \theta^* - \theta \| \tag{9}$$

Proof Denote $y^* := \mathrm{E}[y]$. By (8) and convexity of r,

$$r^* - \mathrm{E}\left[r\left(x^*_\theta, y \right) \right] \leq \mathrm{ME}_\theta \left[r | x^*_\theta \right] - r\left(x^*_\theta, y^* \right)$$

Since r is 1-Lipschitz, this implies

$$r^* - \mathrm{E}\left[r\left(x^*_\theta, y \right) \right] \leq d_Y\!\left(y^*, K_\theta\left(x^*_\theta \right)^+ \right)$$

Using Lemma 21, we get

$$r^* - \mathrm{E}\left[r\left(x^*_\theta, y \right) \right] \leq \left(S^{-1} + 1 \right) d_Y\!\left(y^*, K_\theta\left(x^*_\theta \right)^\flat \right)$$

Applying Lemma 22,

$$r^* - \mathrm{E}\left[r\left(x^*_\theta, y \right) \right] \leq 4 \left(S^{-1} + 1 \right) d_Z(\theta, \mathcal{V}\left(x^*_\theta, y^* \right))$$

It remains to observe that $y^* \in K_{\theta^*}(x^*_\theta)$ and therefore $F(x^*_\theta, \theta^*, y^*) = 0$ and $\theta^* \in \mathcal{V}\left(x^*_\theta, y^* \right)$. ■

We can now start the regret analysis. We consider the probability space resulting from agent policy $\varphi^\eta_{\mathrm{IUCB}}$ and some fixed nature policy. First, we establish that regret can be expressed a sum over cycle contributions, and as long as θ^* is in the confidence set, cycle contributions can be bounded using a certain geometric construction in $\mathcal{Z}$.

We will need some notation. For any $n \in \mathbb{N}$, we denote $[n] := \{ m \in \mathbb{N} \,|\, m < n \}$. For any $N \in \mathbb{N}$ and $X \subseteq [N]$, define the $\mathbb{R}$-valued random variable $\mathrm{Rg}[X]$ by

$$\mathrm{Rg}[X] := \sum_{k \in X} \sum_{n = T_k}^{\min(T_{k+1}, N) - 1} \left(r^* - r\left(x_n, y_n \right) \right)$$

That is, $\mathrm{Rg}[X]$ is the contribution of the cycles in X to regret. For any $a \in \{0, 1\}$ and $k \in \mathbb{N}$, define the $\mathcal{D}$-valued random variable $\bar{y}^a_k$, the $\mathbb{R}_+$-valued random variable ρ^a_k and $\mathbb{N}$-valued random variable τ^a_k by

$$\tau^a_k := \min\left(T_{k+1} - a, N \right) - \min\left(T_k, N \right)$$

$$\bar{y}^a_k := \bar{y}_{T_k, \min(T_{k+1} - a, N)}$$

$$\rho_k^a := \max_{z \in \mathcal{C}_k} d_Z\big(z, \mathcal{V}\left(x_{\theta_k}^*, \bar{y}_k^a\right)\big)$$

That is, τ_k^a is the duration of the intersection of the k-th cycle minus its last a rounds with the time interval of the first N rounds; $\bar{y}_k^a$ is the average of all outcomes in the k-th cycle except for the last a, and also excluding outcomes beyond the time horizon N; and ρ_k^a is the "radius" of the confidence set w.r.t. $\mathcal{V}\left(x_{\theta_k}^*, \bar{y}_k^a\right)$.

Lemma 27 *Let X be a $2^{[N]}$-valued random variable s.t. for all $k \in X$, $\theta^* \in \mathcal{C}_k$. Then,*

$$\mathrm{Rg}\left[X\right] \leq \sum_{k \in X} \min_{a \in \{0,1\}} \big(4\left(S^{-1}+1\right)\tau_k^a \rho_k^a + C \cdot \mathbf{1}_{a=1}\big)$$

Proof Consider any sequence $\{a_k \in \{0,1\}\}_{k<N}$. We have,

$$\mathrm{Rg}\left[X\right] = \sum_{k \in X}\left(\sum_{n=T_k}^{\min(T_{k+1}-a,N)-1}\left(r^* - r\left(x_n, y_n\right)\right) \right.$$
$$\left. + \left(r^* - r\left(x_{T_{k+1}-1}, y_{T_{k+1}-1}\right)\right)\mathbf{1}_{a_k=1}\mathbf{1}_{T_{k+1}\leq N}\right)$$

Hence,

$$\begin{aligned}
\mathrm{Rg}\left[X\right] = & \sum_{k \in X} \min_{a \in \{0,1\}} \left(\sum_{n=T_k}^{\min(T_{k+1}-a,N)-1}\left(r^* - r\left(x_n, y_n\right)\right)\right. \\
& \left. + \left(r^* - r\left(x_{T_{k+1}-1}, y_{T_{k+1}-1}\right)\right)\mathbf{1}_{a=1}\mathbf{1}_{T_{k+1}\leq N}\right) \\
\leq & \sum_{k \in X} \min_{a \in \{0,1\}} \left(\sum_{n=\min(T_k,N)}^{\min(T_{k+1}-a,N)-1}\left(r^* - r\left(x_n, y_n\right)\right) + C \cdot \mathbf{1}_{a=1}\mathbf{1}_{T_{k+1}\leq N}\right) \\
= & \sum_{k \in X} \min_{a \in \{0,1\}} \left(\tau_k^a r^* - \sum_{m<\tau_k^a} r\left(x_{T_k+m}, y_{T_k+m}\right) + C \cdot \mathbf{1}_{a=1}\mathbf{1}_{T_{k+1}\leq N}\right)
\end{aligned}$$

For any $k \in X$, $\theta^* \in \mathcal{C}_k$, and since θ_k are selected optimistically, this implies

$$\mathrm{ME}_{\theta_k}\left[r \middle| x_{\theta_k}^*\right] \geq r^*$$

Hence, we can use Lemma 20 to conclude

$$\mathrm{Rg}\left[X\right] \leq \sum_{k \in X} \min_{a \in \{0,1\}} \left(\tau_k^a d_Y\Big(\bar{y}_k^a, K_{\theta_k}\left(x_{\theta_k}^*\right)^+\Big) + C \cdot \mathbf{1}_{T_{k+1}\leq N}\mathbf{1}_{a=1}\right)$$

By Lemma 21, we have

$$d_Y\left(\bar{y}_k^a, K_{\theta_k}\left(x^*_{\theta_k}\right)\right)^+ \leq \left(S^{-1}+1\right) d_Y\left(\bar{y}_k^a, K_{\theta_k}\left(x^*_{\theta_k}\right)\right)^\flat$$

We get

$$\mathrm{Rg}\left[X\right] \leq \sum_{k\in X} \min_{a\in\{0,1\}} \left(\left(S^{-1}+1\right)\tau_k^a d_Y\left(\bar{y}_k^a, K_{\theta_k}\left(x^*_{\theta_k}\right)^\flat\right) + C\cdot \mathbf{1}_{T_{k+1}\leq N}\mathbf{1}_{a=1}\right)$$

Applying Lemma 22, this becomes

$$\mathrm{Rg}\left[X\right] \leq \sum_{k\in X} \min_{a\in\{0,1\}} \left(4\left(S^{-1}+1\right)\tau_k^a d_Z\left(\theta_k, \mathcal{V}\left(x^*_{\theta_k}, \bar{y}_k^a\right)\right) + C\cdot \mathbf{1}_{T_{k+1}\leq N}\mathbf{1}_{a=1}\right)$$

Since $\theta_k \in \mathcal{C}_k$, we get

$$\mathrm{Rg}\left[X\right] \leq \sum_{k\in X} \min_{a\in\{0,1\}} \left(4\left(S^{-1}+1\right)\tau_k^a \rho_k^a + C\cdot \mathbf{1}_{a=1}\right)$$

■

Next, we split all cycles k in which $\theta^* \in \mathcal{C}_k$ into 3 sets, two of which admit certain geometric constraints and the 3rd has a bounded number of cycles.

Lemma 28 *Let $\lambda > 0$ and $N \in \mathbb{N}$. Define the $2^{[N]}$-valued random variables $\mathrm{A}, \mathrm{B}, \Gamma$ by*

$$\mathrm{A} := \left\{k\in[N] \,\middle|\, T_k < N \textit{ and } \max_{z,z'\in\mathcal{C}_k}\left\|z-z'\right\| \leq 2D_Z\lambda\right\}$$

$$\mathrm{B} := \left\{k\in[N]\setminus\mathrm{A} \,\middle|\, T_k < N \textit{ and } \rho_k^0 \leq 9D_Z^2\lambda\right\}$$

$$\Gamma := \left\{k\in[N]\setminus(\mathrm{A}\cup\mathrm{B}) \mid T_k < N\right\}$$

Then,

$$|\Gamma| \leq \gamma D_Z^2 \ln\frac{D_Z R}{\lambda}$$

Proof Using John's ellipsoid theorem for balanced convex sets[17], we can find an inner product on $\mathcal{Z}$ s.t. for any $z\in\mathcal{Z}$

$$\|z\|_2 \leq \|z\| \leq \sqrt{D_Z}\cdot\|z\|_2 \tag{10}$$

Here, $\|z\|_2$ stands for the norm of z w.r.t. the inner product.

Fix some $\lambda > 0$. We can find some $0 \leq D \leq D_Z$, a non-decreasing sequence $\{J_i \in \mathbb{N}\}_{i\leq D}$ and an orthonormal set $\{e_i \in \mathcal{Z}\}_{i<D}$ s.t.

- For any $0\leq i\leq D$, $0\leq j<i$ and any $\theta,\theta'\in\mathcal{C}_{J_i}$, $|e_j\cdot(\theta-\theta')| \leq 2\lambda$.

17. See Giorgi and Kjeldsen (2013), p. 214.

- For any $0 \leq i < D$, either (a) $J_i = J_{i+1}$ or (b) for any unit vector $v \in \mathcal{Z}$, if for all $j < i$, $e_j \cdot v = 0$, then there are some $\theta, \theta' \in \mathcal{C}_{J_i}$ s.t. $|v \cdot (\theta - \theta')| > 2\lambda$.
- For any $k \in \mathbb{N}$ and unit vector $v \in \mathcal{Z}$, if for all $i < D$, $e_i \cdot v = 0$, then there are some $\theta, \theta' \in \mathcal{C}_k$ s.t. $|v \cdot (\theta - \theta')| > 2\lambda$.

That is, $\{e_i\}$ is constructed by adding vectors iteratively. Once the width of $\mathcal{C}_k$ becomes less than 2λ along any direction orthogonal to the vectors already in the set, we add this direction to the set.

Consider any $0 \leq i \leq D$ and $k \geq J_i$ s.t. $T_k < N$. Choose $\theta^k \in \mathcal{C}_k$ s.t.

$$d_Z\Big(\theta^k, \mathcal{V}\left(x^*_{\theta_k}, \bar{y}_k\right)\Big) = \rho_k^0$$

(Assuming $\mathcal{C}_k \neq \varnothing$, we'll get back to this below.) By the supporting hyperplane theorem, there are $\alpha_k \in \bar{\mathcal{Z}}^\star$ and $c_k \in \mathbb{R}$ s.t.

- $\|\alpha_k\| = 1$
- For all $z \in \mathcal{V}\left(x^*_{\theta_k}, \bar{y}_k\right)$, $\alpha_k(z) = c_k$.
- $\alpha_k(\theta^k) = c_k + \rho_k^0$

For any $\theta \in \mathcal{C}_{k+1}$, we have

$$\begin{aligned}|\alpha_k(\theta) - c_k| &\leq \frac{\eta}{\sqrt{\tau_k}} \\ &\leq \frac{\rho_k^0}{2\left(D_Z + 1\right)}\end{aligned}$$

Here, the first inequality follows from the definition of $\mathcal{C}_{k+1}$ and the fact that $\|\alpha_k\| = 1$, and the second inequality follows from (1).

Let $\beta_k := \alpha_k|_{\mathcal{Z}}$. For all $\theta \in \mathcal{C}_k$:

$$\begin{aligned}|\beta_k(\theta) - c_k| &= |\alpha_k(\theta) - c_k| \\ &\leq \|\alpha_k\| \cdot d_Z\big(\theta, \mathcal{V}\left(x^*_{\theta_k}, \bar{y}_k\right)\big) \\ &\leq \rho_k^0\end{aligned}$$

Moreover,

$$\begin{aligned}\left|\beta_k\left(\theta^k\right) - c_k\right| &= \left|\alpha_k\left(\theta^k\right) - c_k\right| \\ &= \rho_k^0\end{aligned}$$

And, for all $\theta \in \mathcal{C}_{k+1}$:

$$\begin{aligned}|\beta_k(\theta) - c_k| &= |\alpha_k(\theta) - c_k| \\ &\leq \frac{\rho_k^0}{2\left(D_Z + 1\right)}\end{aligned}$$

Observe that

$$\begin{aligned}\|\beta_k\|_2 &\leq \sqrt{D_Z}\cdot\|\beta_k\| \\ &\leq \sqrt{D_Z}\cdot\|\alpha_k\| \\ &= \sqrt{D_Z}\end{aligned}$$

Let $P_i : \mathcal{Z} \to \mathcal{Z}$ be the orthogonal projection to the orthogonal complement of $\{e_j\}_{j<i}$ in $\mathcal{Z}$. We can apply Lemma 25 to a translation of $\mathcal{C}_k$ and $\frac{1}{\sqrt{D_Z}}\beta_k$ to conclude that *one* of the following two inequalities must be true (and, we can drop the assumption $\mathcal{C}_k \neq \varnothing$, because the 1st inequality obviously still holds):

$$\begin{aligned}\rho_k^0 &< \sqrt{D_Z}\cdot\left(4\left(D_Z - i\right) + 5\right)\sqrt{i}\cdot\lambda \\ &\leq 9D_Z^2\lambda\end{aligned}$$

$$\exists c_k' \in \mathbb{R} : \max_{\theta\in\mathcal{C}_{k+1}}\left|\beta_k\left(P_i\theta\right) - c_k'\right| \leq \frac{1}{D_Z - i + 1}\cdot\max_{\theta\in\mathcal{C}_k}\left|\beta_k\left(P_i\theta\right) - c_k'\right| \tag{11}$$

For any $k \in \mathbb{N}$, denote by $\bar{\mathcal{C}}_k$ the convex hull of $\mathcal{C}_k$. For any $0 \leq i < \min(D_Z, D+1)$, denote by $\mathcal{I}_i$ the set of all $J_i \leq k < J_{i+1}$ (for $i = D$, just $k \geq J_i$) for which (11) holds. For any $k \in \mathcal{I}_i$, we can apply Lemma 23 to get

$$vol_i\left(P_i\bar{\mathcal{C}}_{k+1}\right) \leq \left(1 - \frac{1}{e^2}\right)\cdot vol_i\left(P_i\bar{\mathcal{C}}_k\right)$$

Here, vol_i stands for $(D_Z - i)$-dimensional volume inside the orthogonal complement of $\{e_j\}_{j<i}$. For every index in $\mathcal{I}_i$, the volume of $P_i\bar{\mathcal{C}}$ is reduced by a factor of $\Omega(1)$. This volume starts out as $R^{D_Z-i}v_{D_Z-i}$ at most (since $\mathcal{H}$ is contained in a ball of radius R) and, by Lemma 24, cannot go below $\left(\frac{\lambda}{D_Z-i}\right)^{D_Z-i}v_{D_Z-i}$: once the volume is below that, the minimal width is below 2λ which contradicts the condition $k < J_{\mathrm{i}+1}$ (or the definition of D in the case $i = D$). We get that,

$$\begin{aligned}|\mathcal{I}_i| &\leq \frac{1}{\ln\left(1 - \frac{1}{e^2}\right)^{-1}}\cdot\ln\frac{R^{D_Z-i}v_{D_Z-i}}{\left(\frac{\lambda}{D_Z-i}\right)^{D_Z-i}v_{D_Z-i}} \\ &= \gamma\left(D_Z - i\right)\ln\frac{(D_Z-i)R}{\lambda} \\ &\leq \gamma D_Z\ln\frac{D_Z R}{\lambda}\end{aligned}$$

We now see that every k with $T_k < N$ satisfies one of 3 conditions:

1. $\rho_k^0 \leq 9D_Z^2\lambda$
2. $k \in \mathcal{I}_i$ for some $0 \leq i < \min(D_Z, D+1)$
3. $k \geq J_{D_Z}$ (this can only happen for $D = D_Z$)

Consider some k s.t. $T_k < N$ and $\theta^* \in \mathcal{C}_k$. If condition 3 holds, then $\mathcal{C}_k$ is contained in a D_Z-cube with side 2λ and hence $k \in \mathrm{A}$ (notice there's an extra factor of $\sqrt{D_Z}$ in the diameter coming from (10)). If condition 1 holds, then $k \in \mathrm{A} \cup \mathrm{B}$. Therefore, if $k \in \Gamma$ then condition 2 holds. We get

$$\begin{aligned}|\Gamma| \leq & \sum_{i=0}^{\min(D_Z, D-1)} |\mathcal{I}_i| \\ \leq & \gamma D_Z^2 \ln \frac{D_Z R}{\lambda}\end{aligned}$$

■

Let $\Delta := \{k \in [N] \mid \theta^* \in \mathcal{C}_k\}$. Denote $\mathrm{A}_0 := \mathrm{A} \cap \Delta$, $\mathrm{B}_0 := \mathrm{B} \cap \Delta$, $\Gamma_0 := \Gamma \cap \Delta$ and

$$\Delta^c := \{k \in [N] \setminus \Delta \mid T_k < N\}$$

We will establish a high probability bound on $\mathrm{Rg}[\mathrm{A}_0]$.

Lemma 29 *For any $\delta \in (0,1)$, it holds with probability at least $1-\delta$ that*

$$\mathrm{Rg}[\mathrm{A}_0] \leq 8\left(S^{-1}+1\right) D_Z \lambda N + C\sqrt{2N \ln \frac{1}{\delta}}$$

Also,

$$\mathrm{E}[\mathrm{Rg}[\mathrm{A}_0]] \leq 8\left(S^{-1}+1\right) D_Z \lambda N$$

Proof

The expected value in equation (9) of Lemma 26 can be interpreted as the expected value of a the contribution of a particular round to $\mathrm{Rg}[\mathrm{A}_0]$ *conditional* on the past, since the event "$k \in \mathrm{A}_0$, $T_k \leq n < T_{k+1}$" depends only on the past of the n-th round. Hence,

$$\begin{aligned}\mathrm{E}[\mathrm{Rg}[\mathrm{A}_0]] &\leq N \cdot 4\left(S^{-1}+1\right) \cdot 2D_Z \lambda \\ &= 8\left(S^{-1}+1\right) D_Z \lambda N\end{aligned}$$

Here, we use the fact that A_0 only includes cycles in which $\theta^* \in \mathcal{C}_k$, because that's needed for condition (8) and also to bound $\|\theta^* - \theta_k\|$ by the diameter of $\mathcal{C}_k$.

Moreover, the contribution of a single round to regret is in $[-C, +C]$. Hence, by the Azuma-Hoeffding inequality, for any $\Delta > 0$ we have

$$\Pr\left[\mathrm{Rg}[\mathrm{A}_0] \geq 8\left(S^{-1}+1\right) D_Z \lambda N + \Delta\right] \leq \exp\left(-\frac{\Delta^2}{2NC^2}\right)$$

Taking $\Delta := C\sqrt{2N \ln \frac{1}{\delta}}$, we get the desired result.

■

Let's bound the contribution to regret coming from Γ_0.

Lemma 30

$$\mathrm{Rg}\left[\Gamma_0\right] \leq 8\left(S^{-1}+1\right)(D_Z+1)\eta\sqrt{\gamma D_Z^2 \ln\frac{D_Z R}{\lambda}\cdot N} + \gamma C D_Z^2 \ln\frac{D_Z R}{\lambda}$$

Proof Lemma 27 implies (taking $a = 1$)

$$\begin{aligned}\mathrm{Rg}\left[\Gamma_0\right] &\leq \sum_{k\in\Gamma_0}\left(4\left(S^{-1}+1\right)\tau_k^1\rho_k^1 + C\right)\\ &\leq 4\left(S^{-1}+1\right)\sum_{k\in\Gamma_0}\tau_k^1\rho_k^1 + C\left|\Gamma_0\right|\end{aligned}$$

We know that stopping condition (1) doesn't hold before the last round of a cycle, for any $k \in \mathbb{N}$. Hence,

$$\sqrt{\tau_k^1}\cdot\rho_k^1 < 2(D_Z+1)\eta$$

Dividing both sides by $\sqrt{\tau_k^1}$,

$$\rho_k^1 < \frac{2(D_Z+1)\eta}{\sqrt{\tau_k^1}}$$

Using this, we can rewrite the second term on the left hand side of the bound for $\mathrm{Rg}\left[\Gamma_0\right]$:

$$\mathrm{Rg}\left[\Gamma_0\right] \leq 8\left(S^{-1}+1\right)(D_Z+1)\eta\sum_{k\in\Gamma_0}\sqrt{\tau_k^1} + C\left|\Gamma_0\right|$$

Since the sum of the τ_k^1 cannot exceed N, the sum of their square roots cannot exceed $\sqrt{\left|\Gamma_0\right| N}$. We get

$$\mathrm{Rg}\left[\Gamma_0\right] \leq 8\left(S^{-1}+1\right)(D_Z+1)\eta\sqrt{\left|\Gamma_0\right| N} + C\left|\Gamma_0\right|$$

Using Lemma 28, we can bound $\left|\Gamma_0\right| \leq \left|\Gamma\right|$ and get

$$\mathrm{Rg}\left[\Gamma_0\right] \leq 8\left(S^{-1}+1\right)(D_Z+1)\eta\sqrt{\gamma D_Z^2 \ln\frac{D_Z R}{\lambda}\cdot N} + \gamma C D_Z^2 \ln\frac{D_Z R}{\lambda}$$

■

To complete the proof of the main theorem, it remains only to bound the contributions of B_0 and Δ^c. Denote $\mathrm{Rg} := \mathrm{Rg}[[N]]$ (the total regret).

Proof [Proof of Theorem 2]

We have

$$\mathrm{Rg} = \mathrm{Rg}\left[\mathrm{A}_0\right] + \mathrm{Rg}\left[\mathrm{B}_0\right] + \mathrm{Rg}\left[\Gamma_0\right] + \mathrm{Rg}\left[\Delta^c\right]$$

For B_0, Lemma 27 implies (taking $a = 0$)

$$\begin{aligned}\mathrm{Rg}\left[\mathrm{B}_0\right] &\leq \sum_{k<N} 4\left(S^{-1}+1\right)\tau_k^0 \cdot 9D_Z^2\lambda \\ &= 36\left(S^{-1}+1\right)D_Z^2\lambda N\end{aligned}$$

For Δ^c, by Lemma 19,

$$\Pr\left[\Delta^c \neq \varnothing\right] \leq D_W N(N+1)\exp\left(-\Omega(1)\cdot\frac{\eta^2}{R^2 D_W^{\frac{5}{3}}}\right)$$

Since $\mathrm{Rg}\left[\varnothing\right] = 0$ and $\mathrm{Rg}\left[\Delta^c\right] \leq CN$, this implies

$$\begin{aligned}\mathrm{E}\left[\mathrm{Rg}\left[\Delta^c\right]\right] &\leq D_W N(N+1)\exp\left(-\Omega(1)\cdot\frac{\eta^2}{R^2 D_W^{\frac{5}{3}}}\right)\cdot CN \\ &= CD_W N^2(N+1)\exp\left(-\Omega(1)\cdot\frac{\eta^2}{R^2 D_W^{\frac{5}{3}}}\right)\end{aligned}$$

Combining this with Lemma 29 and Lemma 30, we get,

$$\begin{aligned}\mathrm{E}\left[\mathrm{Rg}\right] \leq\ & 8\left(S^{-1}+1\right)D_Z\lambda N \\ & +36\left(S^{-1}+1\right)D_Z^2\lambda N \\ & +8\eta\left(S^{-1}+1\right)(D_Z+1)\sqrt{\gamma D_Z^2 \ln\frac{D_Z R}{\lambda}\cdot N} \\ & +\gamma C D_Z^2 \ln\frac{D_Z R}{\lambda} \\ & +CD_W N^2(N+1)\exp\left(-\Omega(1)\cdot\frac{\eta^2}{R^2 D_W^{\frac{5}{3}}}\right) \\ =\ & 8\eta\left(S^{-1}+1\right)D_Z(D_Z+1)\sqrt{\gamma\ln\frac{D_Z R}{\lambda}\cdot N} \\ & +CD_W N^2(N+1)\exp\left(-\Omega(1)\cdot\frac{\eta^2}{R^2 D_W^{\frac{5}{3}}}\right) \\ & +\gamma C D_Z^2 \ln\frac{D_Z R}{\lambda} \\ & +\left(S^{-1}+1\right)D_Z\left(36D_Z+8\right)N\lambda\end{aligned}$$

Since this holds for any nature policy, this is a bound on $\mathrm{ERg}_{\theta^*}\left(\varphi^\eta_{\mathrm{IUCB}}; N\right)$. ■

Theorem 3 is proved similarly.

Proof [Proof of Theorem 3]

As in the proof of Theorem 2, Lemma 27 implies

$$\mathrm{Rg}\left[\mathrm{B}_0\right] \leq 36\left(S^{-1}+1\right) D_Z^2 \lambda N$$

Also, by Lemma 30

$$\mathrm{Rg}\left[\Gamma_0\right] \leq 8\left(S^{-1}+1\right)(D_Z+1)\eta\sqrt{\gamma D_Z^2 \ln\frac{D_Z R}{\lambda}\cdot N}+\gamma C D_Z^2 \ln\frac{D_Z R}{\lambda}$$

By Lemma 29, with probability at least $1-\delta$,

$$\mathrm{Rg}\left[\mathrm{A}_0\right] \leq 8\left(S^{-1}+1\right) D_Z \lambda N + C\sqrt{2N\ln\frac{1}{\delta}}$$

By Lemma 19, with probability at least $1 - D_W N(N+1)\exp(-\Omega(1)\cdot\frac{\eta^2}{R^2 D_W^{\frac{5}{3}}})$,

$$\mathrm{Rg}\left[\Delta^c\right] = 0$$

Putting everything together, and using a union bound, we get that with probability at least $1-\tilde{\delta}$

$$\begin{aligned}
\mathrm{Rg} \leq \quad & 8\left(S^{-1}+1\right) D_Z \lambda N \\
& + C\sqrt{2N\ln\frac{1}{\delta}} \\
& + 36\left(S^{-1}+1\right) D_Z^2 \lambda N \\
& + 8\left(S^{-1}+1\right)(D_Z+1)\eta\sqrt{\gamma D_Z^2 \ln\frac{D_Z R}{\lambda}\cdot N} \\
& + \gamma C D_Z^2 \ln\frac{D_Z R}{\lambda} \\
= \quad & 8\eta\left(S^{-1}+1\right) D_Z(D_Z+1)\sqrt{\gamma \ln\frac{D_Z R}{\lambda}\cdot N} \\
& + C\sqrt{2N\ln\frac{1}{\delta}} \\
& + C D_W N^2 (N+1)\exp\left(-\Omega(1)\cdot\frac{\eta^2}{R^2 D_W^{\frac{5}{3}}}\right) \\
& + \gamma C D_Z^2 \ln\frac{D_Z R}{\lambda} \\
& + \left(S^{-1}+1\right) D_Z\left(36 D_Z + 8\right) N\lambda
\end{aligned}$$

Since this holds for any nature policy, this is a bound on $\mathrm{Rg}_\theta(\varphi_{\mathrm{IUCB}}; N, \tilde{\delta})$. ■

Appendix B. Some Bounds on R

B.1 K(x) of Codimension 1

We start with a simple fact about norms on dual spaces.

Lemma 31 *Let $\mathcal{X}$ be a finite-dimensional normed space. Then, for any $x \in \mathcal{X}$:*

$$\min_{\alpha \in \mathcal{X}^\star : \alpha(x)=1} \|\alpha\| = \|x\|^{-1}$$

Proof Let $\alpha \in \mathcal{X}^\star$ be any s.t. $\alpha(x) = 1$. Then, $\|\alpha\| \cdot \|x\| \geq 1$, implying that $\|\alpha\| \geq \|x\|^{-1}$ and hence

$$\min_{\alpha \in \mathcal{X}^\star : \alpha(x)=1} \|\alpha\| \geq \|x\|^{-1}$$

On the other hand, consider $\mathcal{X}_0 := \mathbb{R}x$. Consider the unique $\alpha_0 \in \mathcal{X}_0^\star$ s.t. $\alpha_0(x) = 1$. Then, $\|\alpha_0\| = \|x\|^{-1}$. By the Hanh-Banach theorem, this implies there is $\alpha \in \mathcal{X}^\star$ s.t. $\|\alpha\| = \|x\|^{-1}$ and $\alpha|_{\mathcal{X}_0} = \alpha_0$, the latter implying that $\alpha(x) = 1$. Therefore,

$$\min_{\alpha \in \mathcal{X}^\star : \alpha(x)=1} \|\alpha\| \leq \|x\|^{-1}$$

■

The norm of an operator bounds how much applying it can increase the norm of a vector. Similarly, the norm of the inverse operator bounds how much applying the original operator can *decrease* the norm of a vector.

Lemma 32 *Let $\mathcal{X}$ and $\mathcal{X}'$ be finite-dimensional normed spaces, $x \in \mathcal{X}$ and $A : \mathcal{X} \to \mathcal{X}'$ an invertible operator. Then,*

$$\|Ax\| \geq \frac{\|x\|}{\|A^{-1}\|}$$

Proof We have,

$$\begin{aligned} \|x\| &= \left\|A^{-1}Ax\right\| \\ &\leq \left\|A^{-1}\right\| \cdot \|Ax\| \end{aligned}$$

Dividing both sides of the inequality by $\left\|A^{-1}\right\|$ we get the desired result. ■

Proposition 4 now follows from a straightforward calculation.

Proof [Proof of Proposition 4]First, let's examine the standard norm $\|\cdot\|_{\mathcal{W}}$ on $\mathcal{W}$. Since $\mathcal{W} \cong \mathbb{R}$, it is enough to compute $\|1\|_{\mathcal{W}}$:

$$
\begin{aligned}
\|1\|_{\mathcal{W}} &= \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \min_{y \in \mathcal{Y}: F_{x\theta} y = 1} \|y\| \\
&= \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \min_{y \in \mathcal{Y}: (A_x \theta) y = 1} \|y\| \\
&= \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \|A_x \theta\|^{-1} \\
&= \frac{1}{\min_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \|A_x \theta\|}
\end{aligned}
$$

Here, we used the definition of A_x followed by Lemma 31 applied to $\mathcal{X} := \mathcal{Y}^\star$.

Second, let's examine the standard norm on $\mathcal{Z}$. We have,

$$
\begin{aligned}
\|z\| &= \max_{x \in \mathcal{A}} \|F_{xz}\| \\
&= \max_{x \in \mathcal{A}} \left(\|A_x z\| \cdot \|1\|_{\mathcal{W}} \right) \\
&= \frac{\max_{x \in \mathcal{A}} \|A_x z\|}{\min_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \|A_x \theta\|}
\end{aligned}
$$

Here, we observed that the induced norm on the space of operators from $\mathcal{Y}$ to $\mathcal{W}$ is the same as the induced norm on $\mathcal{Y}^\star$ multiplied by the scalar $\|1\|_{\mathcal{W}}$.

Finally, we get,

$$
\begin{aligned}
R(\mathcal{H}, F) &= \max_{\theta \in \mathcal{H}} \|\theta\| \\
&= \frac{\max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \|A_x \theta\|}{\min_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \|A_x \theta\|} \\
&\leq \frac{\max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \left(\|A_x\|_0 \cdot \|\theta\|_0 \right)}{\min_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \left(\left\|A_x^{-1}\right\|_0^{-1} \cdot \|\theta\|_0 \right)} \\
&\leq \frac{\left(\max_{x \in \mathcal{A}} \|A_x\|_0\right) \left(\max_{\theta \in \mathcal{H}} \|\theta\|_0\right)}{\left(\min_{x \in \mathcal{A}} \left\|A_x^{-1}\right\|_0^{-1}\right) \left(\min_{\theta \in \mathcal{H}} \|\theta\|_0\right)} \\
&= \frac{\max_{\theta \in \mathcal{H}} \|\theta\|_0}{\min_{\theta \in \mathcal{H}} \|\theta\|_0} \left(\max_{x \in \mathcal{A}} \|A_x\|_0 \right) \left(\max_{a \in \mathcal{A}} \|A_x^{-1}\|_0 \right)
\end{aligned}
$$

Here we used Lemma 32 on the third line. ■

Proposition 5 follows easily from Proposition 4.

Proof [Proof of Proposition 5]Set χ and λ to be

$$
\chi(\theta) := \frac{1}{\|\theta\|_0}
$$

$$\lambda(x) := \left\| A_x^{-1} \right\|_0$$

Notice that the bandit specified by $\mathcal{H}^\chi$ and F^λ is of a form to which Proposition 4 is applicable, with the operators A_x replaced by $\tilde{A}_x := \lambda(x) A_x$. Applying Proposition 4 we get

$$\begin{aligned}
R\left(\mathcal{H}^\chi, F^\lambda\right) &= \frac{\max_{\theta \in \mathcal{H}^\chi} \|\theta\|_0}{\min_{\theta \in \mathcal{H}^\chi} \|\theta\|_0} \left(\max_{x \in \mathcal{A}} \|\tilde{A}_x\|_0 \right) \left(\max_{a \in \mathcal{A}} \|\tilde{A}_x^{-1}\|_0 \right) \\
&= \frac{\max_{\theta \in \mathcal{H}} \chi(\theta) \|\theta\|_0}{\min_{\theta \in \mathcal{H}} \chi(\theta) \|\theta\|_0} \left(\max_{x \in \mathcal{A}} \lambda(x) \|A_x\|_0 \right) \left(\max_{a \in \mathcal{A}} \lambda(x)^{-1} \|A_x^{-1}\|_0 \right) \\
&= 1 \cdot \left(\max_{x \in \mathcal{A}} \left\| A_x^{-1} \right\|_0 \cdot \|A_x\|_0 \right) \left(\max_{a \in \mathcal{A}} \left\| A_x^{-1} \right\|_0^{-1} \cdot \|A_x^{-1}\|_0 \right) \\
&= \max_{x \in \mathcal{A}} \left\| A_x^{-1} \right\|_0 \cdot \|A_x\|_0
\end{aligned}$$

■

B.2 K(x) of Dimension 1

Proof [Proof of Proposition 6]Since $\mathcal{W}$ is a subspace of $\mathcal{Y}$, it is equipped with two norms: the standard norm $\|\cdot\|_{\mathcal{W}}$ on $\mathcal{W}$ and the norm $\|\cdot\|_{\mathcal{Y}}$ induced by the norm on $\mathcal{Y}$. Observe that for any $x \in \mathcal{A}$, $\theta \in \mathcal{H}$ and $w \in \mathcal{W}$,

$$\begin{aligned}
F_{x\theta} w &= \psi(\theta) w - \mu(w) f(x, \theta) \\
&= 1 \cdot w - 0 \cdot f(x, \theta) \\
&= w
\end{aligned}$$

It follows that

$$\begin{aligned}
\|w\|_{\mathcal{W}} &= \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \min_{y \in \mathcal{Y} : F_{x\theta} y = w} \|y\|_{\mathcal{Y}} \\
&\leq \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \|w\|_{\mathcal{Y}} \\
&= \|w\|_{\mathcal{Y}}
\end{aligned}$$

We get

$$
\begin{aligned}
R(\mathcal{H}, F) &= \max_{\theta \in \mathcal{H}} \|\theta\| \\
&= \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \|F_{x\theta}\| \\
&= \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \max_{y \in \mathcal{Y}: \|y\|_{\mathcal{Y}} = 1} \|F_{x\theta} y\|_{\mathcal{W}} \\
&\leq \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \max_{y \in \mathcal{Y}: \|y\|_{\mathcal{Y}} = 1} \|F_{x\theta} y\|_{\mathcal{Y}} \\
&= \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \max_{y \in \mathcal{Y}: \|y\|_{\mathcal{Y}} = 1} \|\psi(\theta) y - \mu(y) f(x, \theta)\|_{\mathcal{Y}} \\
&= \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \max_{y \in \mathcal{Y}: \|y\|_{\mathcal{Y}} = 1} \|y - \mu(y) f(x, \theta)\|_{\mathcal{Y}} \\
&\leq \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \max_{y \in \mathcal{Y}: \|y\|_{\mathcal{Y}} = 1} \left(\|y\|_{\mathcal{Y}} + |\mu(y)| \cdot \|f(x, \theta)\|_{\mathcal{Y}}\right) \\
&= \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \max_{y \in \mathcal{Y}: \|y\|_{\mathcal{Y}} = 1} \left(\|y\|_{\mathcal{Y}} + |\mu(y)|\right) \\
&\leq \max_{\substack{x \in \mathcal{A} \\ \theta \in \mathcal{H}}} \max_{y \in \mathcal{Y}: \|y\|_{\mathcal{Y}} = 1} \left(\|y\|_{\mathcal{Y}} + \|y\|_{\mathcal{Y}}\right) \\
&= 2
\end{aligned}
$$

■

Appendix C. Some Bounds on S

C.1 Sine with Simplex

We now proceed to proving the propositions in subsection 5.2.

The following lemma implies that when $\mathcal{D} = \Delta\mathcal{B}$, the norm on $\mathcal{Y} = \mathbb{R}^{\mathcal{B}}$ is the ℓ_1 norm.

Lemma 33 *For any finite set $\mathcal{B}$, the absolute convex hull of $\Delta\mathcal{B}$ in $\mathbb{R}^{\mathcal{B}}$ is the unit ball of the ℓ_1 norm.*

Proof Let y be in the absolute convex hull of $\Delta\mathcal{B}$. Then, there are $n \in \mathbb{N}$, $\{c_i \in \mathbb{R}\}_{i<n}$ and $\{y_i \in \Delta\mathcal{B}\}_{i<n}$ s.t.

$$\sum_{i=0}^{n-1} |c_i| \leq 1$$

$$y = \sum_{i=0}^{n-1} c_i y_i$$

It follows that

$$\begin{aligned}\|y\|_1 &= \left\| \sum_{i=0}^{n-1} c_i y_i \right\|_1 \\ &\leq \sum_{i=0}^{n-1} \|c_i y_i\|_1 \\ &= \sum_{i=0}^{n-1} |c_i| \cdot \|y_i\|_1 \\ &= \sum_{i=0}^{n-1} |c_i| \\ &\leq 1\end{aligned}$$

Here, we used that $\|y_i\| = 1$ since $y_i \in \Delta\mathcal{B}$.
Conversly, let $y \in \mathbb{R}^{\mathcal{B}}$ be s.t. $\|y\|_1 \leq 1$. Let $y^-, y^+ \in \Delta\mathcal{B}$ be s.t. for any $a \in \mathcal{B}$

$$y_a^- := \frac{\min(y_a, 0)}{\sum_{b \in \mathcal{B}} \min(y_b, 0)}$$

$$y_a^+ := \frac{\max(y_a, 0)}{\sum_{b \in \mathcal{B}} \max(y_b, 0)}$$

Here, if the denominator vanishes then we choose any element of $\Delta\mathcal{B}$ arbitrarily. We get,

$$\begin{aligned}y_a &= \min(y_a, 0) + \max(y_a, 0) \\ &= \sum_{b \in \mathcal{B}} \min(y_b, 0)\, y_a^- + \sum_{b \in \mathcal{B}} \max(y_b, 0)\, y_a^+\end{aligned}$$

Hence,

$$y = \sum_{b \in \mathcal{B}} \min(y_b, 0) \cdot y^- + \sum_{b \in \mathcal{B}} \max(y_b, 0) \cdot y^+$$

Moreover,

$$\begin{aligned}\left| \sum_{b \in \mathcal{B}} \min(y_b, 0) \right| + \left| \sum_{b \in \mathcal{B}} \max(y_b, 0) \right| &\leq \sum_{b \in \mathcal{B}} |\min(y_b, 0)| + \sum_{b \in \mathcal{B}} |\max(y_b, 0)| \\ &= \sum_{b \in \mathcal{B}} (|\min(y_b, 0)| + |\max(y_b, 0)|) \\ &= \sum_{b \in \mathcal{B}} |y_b| \\ &= \|y\|_1 \\ &\leq 1\end{aligned}$$

Therefore, y is in the absolute convex hull of $\Delta\mathcal{B}$. ■

In order to analyze sines in the case $\mathcal{D} = \Delta\mathcal{B}$, we will need an expression for the ℓ_1 distance between the simplex $\Delta\mathcal{B}$ and points in its affine hull.

Lemma 34 *For any finite set $\mathcal{B}$ and $y \in \mathbb{R}^B$ s.t. $\sum_a y_a = 1$,*

$$d_1(y, \Delta\mathcal{B}) = \sum_{a \in \mathcal{B}} |y_a| - 1$$

where d_1 stands for minimal ℓ_1 distance.

Proof Define $\mathcal{B}_\pm$ by

$$\mathcal{B}_+ := \{a \in \mathcal{B} \mid y_a \geq 0\}$$

$$\mathcal{B}_- := \{a \in \mathcal{B} \mid y_a < 0\}$$

For any $y' \in \Delta\mathcal{B}$, we have

$$\begin{aligned}
\left\|y - y'\right\|_1 &= \sum_{a \in \mathcal{B}} \left|y_a - y'_a\right| \\
&= \sum_{a \in \mathcal{B}_+} \left|y_a - y'_a\right| + \sum_{a \in \mathcal{B}_-} \left|y_a - y'_a\right| \\
&\geq \left|\sum_{a \in \mathcal{B}_+} \left(y_a - y'_a\right)\right| + \sum_{a \in \mathcal{B}_-} \left(y'_a - y_a\right) \\
&= \left|\sum_{a \in \mathcal{B}_+} y_a - \sum_{a \in \mathcal{B}_+} y'_a\right| + \sum_{a \in \mathcal{B}_-} y'_a - \sum_{a \in \mathcal{B}_-} y_a \\
&= \sum_{a \in \mathcal{B}_+} |y_a| - \sum_{a \in \mathcal{B}_+} y'_a + \sum_{a \in \mathcal{B}_-} y'_a + \sum_{a \in \mathcal{B}_-} |y_a| \\
&= \sum_{a \in \mathcal{B}} |y_a| - \sum_{a \in \mathcal{B}_+} y'_a + \sum_{a \in \mathcal{B}_-} y'_a \\
&\geq \sum_{a \in \mathcal{B}} |y_a| - \sum_{a \in \mathcal{B}_+} y'_a - \sum_{a \in \mathcal{B}_-} y'_a \\
&= \sum_{a \in \mathcal{B}} |y_a| - 1
\end{aligned}$$

Here, we used the fact that $\sum_{a \in \mathcal{B}_+} y_a \geq 1 \geq \sum_{a \in \mathcal{B}_+} y'_a$ on the 5th line.
Conversly, define $y^* \in \Delta\mathcal{B}$ by

$$y^*_a := \frac{\max(y_a, 0)}{\sum_{b \in \mathcal{B}} \max(y_b, 0)}$$

We have,

$$\begin{aligned}
\|y - y^*\|_1 &= \sum_{a\in\mathcal{B}} \left| y_a - \frac{\max(y_a, 0)}{\sum_{b\in\mathcal{B}} \max(y_b, 0)} \right| \\
&= \sum_{a\in\mathcal{B}_+} \left| y_a - \frac{\max(y_a, 0)}{\sum_{b\in\mathcal{B}} \max(y_b, 0)} \right| + \sum_{a\in\mathcal{B}_-} \left| y_a - \frac{\max(y_a, 0)}{\sum_{b\in\mathcal{B}} \max(y_b, 0)} \right| \\
&= \sum_{a\in\mathcal{B}_+} \left| y_a - \frac{y_a}{\sum_{b\in\mathcal{B}_+} y_b} \right| + \sum_{a\in\mathcal{B}_-} |y_a| \\
&= \left(1 - \frac{1}{\sum_{b\in\mathcal{B}_+} y_b}\right) \sum_{a\in\mathcal{B}_+} y_a + \sum_{a\in\mathcal{B}_-} |y_a| \\
&= \sum_{a\in\mathcal{B}_+} y_a - 1 + \sum_{a\in\mathcal{B}_-} |y_a| \\
&= \sum_{a\in\mathcal{B}} |y_a| - 1
\end{aligned}$$

■

Since Proposition 7 is a special case of Proposition 8 (for $\mathcal{E} = \mathcal{B}$), it is sufficient to prove the latter. The proof proceeds by connecting an arbitrary point in $y \in \mathcal{U}^\flat \setminus \Delta\mathcal{B}$ by a straight line segment with the point in $\mathcal{U}^+$ in which the maximum on the right hand side is achieved. The point where this segment meets the boundary of the simplex is used to produce an upper bound on the distance from y to $\mathcal{U}^+$.

Proof [Proof of Proposition 8]Define y^* by

$$y^* := \operatorname*{argmax}_{y\in\mathcal{U}^+} \min_{a\in\mathcal{E}} y_a$$

Consider some $y \in \mathcal{U}^\flat \setminus \Delta\mathcal{B}$. Define a^* by

$$a^* := \operatorname*{argmax}_{a\in\mathcal{B}:y_a<0} \frac{|y_a|}{|y_a| + y^*_a}$$

Notice that $a^* \in \mathcal{E}$ since for any $a \notin \mathcal{E}$, $y_a = 0$. Define t^* by

$$\begin{aligned}
t^* &:= \max_{a\in\mathcal{B}:y_a<0} \frac{|y_a|}{|y_a| + y^*_a} \\
&= \frac{|y_{a^*}|}{|y_{a^*}| + y^*_{a^*}}
\end{aligned}$$

And, define y' by

$$y' := t^* y^* + (1 - t^*)\, y$$

Finally, define $\mathcal{B}_\pm$ by

$$\mathcal{B}_+ := \{a \in \mathcal{B} \mid y_a \geq 0\}$$

$$\mathcal{B}_- := \{a \in \mathcal{B} \mid y_a < 0\}$$

Then, for any $a \in \mathcal{B}_+$

$$\begin{aligned} y'_a &= t^* y^*_a + (1 - t^*)\, y_a \\ &\geq 0 \end{aligned}$$

And, for any $a \in \mathcal{B}_-$

$$\begin{aligned} y'_a &= t^* y^*_a + (1 - t^*)\, y_a \\ &= y_a + t^* \left(y^*_a - y_a\right) \\ &= y_a + t^* \left(|y_a| + y^*_a\right) \\ &\geq y_a + \frac{|y_a|}{|y_a| + y^*_a} \cdot \left(|y_a| + y^*_a\right) \\ &= y_a + |y_a| \\ &= 0 \end{aligned}$$

We conclude $y' \in \Delta\mathcal{B}$. Since it's also clear that $y' \in \mathcal{U}^\flat$, we get $y' \in \mathcal{U}^+$. Hence,

$$\begin{aligned} d_1\left(y, \mathcal{U}^+\right) &\leq \left\|y' - y\right\|_1 \\ &= \left\|\frac{t^*}{1 - t^*} \cdot \left(y^* - y'\right)\right\|_1 \\ &= \frac{t^*}{1 - t^*} \cdot \left\|y^* - y'\right\|_1 \\ &\leq \frac{t^*}{1 - t^*} \cdot \left(\|y^*\|_1 + \left\|y'\right\|_1\right) \\ &= \frac{2t^*}{1 - t^*} \\ &= \frac{2\,|y_{a^*}|}{y^*_{a^*}} \end{aligned}$$

Moreover, by Lemma 34

$$\begin{aligned}
d_1(y, \Delta\mathcal{B}) &= \sum_{a\in\mathcal{B}} |y_a| - 1 \\
&= |y_{a^*}| + \sum_{a\in\mathcal{B}\setminus a^*} |y_a| - 1 \\
&\geq |y_{a^*}| + \sum_{a\in\mathcal{B}\setminus a^*} y_a - 1 \\
&= |y_{a^*}| + (1 - y_{a^*}) - 1 \\
&= 2\,|y_{a^*}|
\end{aligned}$$

Combining the last two inequalites, we get

$$\begin{aligned}
\frac{d_1(y, \Delta\mathcal{B})}{d_1(y, \mathcal{U}^+)} &\geq 2\,|y_{a^*}| \cdot \frac{y^*_{a^*}}{2\,|y_{a^*}|} \\
&= y^*_{a^*} \\
&\geq \min_{a\in\mathcal{E}} y^*_a \\
&= \max_{y\in\mathcal{U}^+} \min_{a\in\mathcal{E}} y_a
\end{aligned}$$

By Lemma 33, d_1 is the same thing as d_Y. And, since this holds for *any* $y \in \mathcal{U}^\flat \setminus \Delta\mathcal{B}$, we conclude

$$\sin\left(\mathcal{U}^\flat, \Delta\mathcal{B}\right) \geq \max_{y\in\mathcal{U}^+} \min_{i\in\mathcal{E}} y_i$$

■

C.2 Sine with Ball

The following elementary inequality will be needed for the proof of Proposition 9.

Lemma 35 *The following inequality holds for any $\alpha \in [0, \frac{\pi}{2}]$ and $t > \sin\alpha$:*

$$\frac{\sqrt{t^2 + \cos^2\alpha} - 1}{t - \sin\alpha} \geq \sin\alpha$$

Proof We have

$$\begin{aligned}
t^2 - 2(\sin\alpha)\,t + 1 - \cos^2\alpha &= t^2 - 2(\sin\alpha)\,t + \sin^2\alpha \\
&= (t - \sin\alpha)^2 \\
&> 0
\end{aligned}$$

Multiplying both sides by $\cos^2\alpha$ we get

$$\left(\cos^2\alpha\right)t^2 - 2\left(\cos^2\alpha\right)\left(\sin\alpha\right)t + \cos^2\alpha - \cos^4\alpha \geq 0$$

Applying the identity $\cos^2\alpha = 1 - \sin^2\alpha$ to the first term on the left hand side, we get

$$t^2 - (\sin^2\alpha)t^2 - 2\left(\cos^2\alpha\right)\left(\sin\alpha\right)t + \cos^2\alpha - \cos^4\alpha \geq 0$$

Now we move all the negative terms to the right hand side:

$$\begin{aligned} t^2 + \cos^2\alpha &\geq \left(\sin^2\alpha\right)t^2 + 2\left(\cos^2\alpha\right)\left(\sin\alpha\right)t + \cos^4\alpha \\ &= \left(\left(\sin\alpha\right)t + \cos^2\alpha\right)^2 \end{aligned}$$

Taking square root of both sides, we get

$$\begin{aligned} \sqrt{t^2 + \cos^2\alpha} &\geq \left(\sin\alpha\right)t + \cos^2\alpha \\ &= \left(\sin\alpha\right)t + 1 - \sin^2\alpha \end{aligned}$$

Moving 1 to the left hand side we get

$$\sqrt{t^2 + \cos^2\alpha} - 1 \geq \left(\sin\alpha\right)\left(t - \sin\alpha\right)$$

Finally, we divide both sides by $t - \sin\alpha$ and conclude that

$$\frac{\sqrt{t^2 + \cos^2\alpha} - 1}{t - \sin\alpha} \geq \sin\alpha$$

■

When $\mathcal{D}$ is a ball, the induced metric on its affine hull is Euclidean. For any $y \in \mathcal{U}^\flat \setminus \mathcal{D}$, we connect y by a straight line segment to the point of $\mathcal{U}^\flat$ nearest to the center of the ball. The point where this segment meets the sphere is the nearest point to y inside $\mathcal{U}^+$.

Proof [Proof of Proposition 9]Define $\mathcal{U}_0$ by

$$\mathcal{U}_0 := \left\{ y \in \mathbb{R}^n \;\middle|\; \begin{bmatrix} y \\ 1 \end{bmatrix} \in \mathcal{U} \right\}$$

We also denote by $\mathcal{D}_0$ the closed unit ball in $\mathbb{R}^n$. Define y^* and ρ by

$$y^* := \underset{y \in \mathcal{U}_0}{\operatorname{argmin}} \|y\|_2$$

$$\begin{aligned} \rho &:= \min_{y \in \mathcal{U}_0} \|y\|_2 \\ &= \|y^*\|_2 \end{aligned}$$

Consider some $y \in \mathcal{U}_0$ s.t. $\|y\|_2 > 1$. Define t by

$$t := \|y - y^*\|_2$$

Denoting minimal ℓ_2 distance by d_2, we have

$$\begin{aligned} d_2(y, \mathcal{D}_0) &= \|y\|_2 - 1 \\ &= \|y^* + (y - y^*)\|_2 - 1 \\ &= \sqrt{\|y^*\|_2^2 + \|y - y^*\|_2^2 + 2(y - y^*)^{\mathrm{t}} y^*} - 1 \\ &= \sqrt{\rho^2 + t^2} - 1 \end{aligned}$$

Here, we observed that $(y - y^*)^{\mathrm{t}} y^* = 0$ due to the definition of y^*.

Let y' be a point on the straight line segment $y^* y$ at distance $\sqrt{1-\rho^2}$ from y^*. Then, $\|y'\| = 1$, and hence $y' \in \mathcal{U}_0 \cap \mathcal{D}_0$ and

$$\begin{aligned} d_2(y, \mathcal{U}_0 \cap \mathcal{D}_0) &\leq \|y - y'\|_2 \\ &= t - \sqrt{1-\rho^2} \end{aligned}$$

We get,

$$\begin{aligned} \frac{d_2(y, \mathcal{D}_0)}{d_2(y, \mathcal{U}_0 \cap \mathcal{D}_0)} &\geq \frac{\sqrt{t^2 + \rho^2} - 1}{t - \sqrt{1-\rho^2}} \\ &\geq \sqrt{1-\rho^2} \end{aligned}$$

Here, we used Lemma 35 with $\alpha = \arccos \rho$ on the second line.

By rotational symmetry, the metric on $\mu^{-1}(1)$ is Euclidean up to a scalar. Hence, we can conclude

$$\sin\left(\mathcal{U}^\flat, \mathcal{D}\right) \geq \sqrt{1-\rho^2}$$

■

Appendix D. Probability Systems (Proofs)

D.1 Single Conditional Probability

We prove Proposition 11 first and Proposition 10 second, because the latter proof will use a special case of the former proposition.

In order to prove Proposition 11, we will need the following lemma. For any point y in the affine hull of the simplex $\Delta\mathcal{B}$, it characterizes some (actually all) points in $\Delta\mathcal{B}$ that are at minimal ℓ_1 distance from y.

Lemma 36 *Let $\mathcal{B}$ be a finite set, $y \in \mathbb{R}^{\mathcal{B}}$ s.t. $\sum_a y_a = 1$ and $\xi \in \Delta\mathcal{B}$. Assume that for any $a \in \mathcal{B}$:*

- *If* $y_a < 0$ *then* $\xi_a = 0$.
- *If* $y_a \geq 0$ *then* $\xi_a \leq y_a$.

Then,

$$\|y - \xi\|_1 = d_1(y, \Delta\mathcal{B})$$

Proof We have,

$$\begin{aligned}
\|y - \xi\|_1 &= \sum_{a \in \mathcal{B}} |y_a - \xi_a| \\
&= \sum_{a \in \mathcal{B}: y_a < 0} |y_a - \xi_a| + \sum_{a \in \mathcal{B}: y_a \geq 0} |y_a - \xi_a| \\
&= \sum_{a \in \mathcal{B}: y_a < 0} |y_a| + \sum_{a \in \mathcal{B}: y_a \geq 0} (y_a - \xi_a) \\
&= \sum_{a \in \mathcal{B}: y_a < 0} |y_a| + \sum_{a \in \mathcal{B}: y_a \geq 0} |y_a| - \sum_{a \in \mathcal{B}: y_a \geq 0} \xi_a \\
&= \sum_{a \in \mathcal{B}} |y_a| - \sum_{a \in \mathcal{B}: y_a \geq 0} \xi_a - \sum_{a \in \mathcal{B}: y_a < 0} \xi_a \\
&= \sum_{a \in \mathcal{B}} |y_a| - \sum_{a \in \mathcal{B}} \xi_a \\
&= \sum_{a \in \mathcal{B}} |y_a| - 1 \\
&= d_1(y, \Delta\mathcal{B})
\end{aligned}$$

Here, we used Lemma 34 on the last line. ■

The proof of Proposition 11 works by constructing a point in $\Delta\mathcal{B}$ that satisfies the conditions of the previous lemma, while staying inside $\mathcal{U}$.

Proof [Proof of Proposition 11]Denote $\mathcal{E}_+ := \mathcal{E} \cap \mathcal{E}'$, $\mathcal{E}_- := \mathcal{E} \setminus \mathcal{E}'$, $\mathcal{E}_0 := \mathcal{B} \setminus \mathcal{E}$. For any $y \in \mathbb{R}^{\mathcal{B}}$, $y \in \mathcal{U}$ if and only if

$$(1 - p) \sum_{a \in \mathcal{E}_+} y_a = p \sum_{a \in \mathcal{E}_-} y_a \tag{12}$$

Consider any $y \in \mathcal{U}^\flat \setminus \Delta\mathcal{B}$. By Lemma 33, it's enough to prove that

$$d_1\left(y, \mathcal{U}^+\right) = d_1(y, \Delta\mathcal{B})$$

We split the proof into 3 cases. The first case is when

$$\sum_{a \in \mathcal{E}_+} y_a \leq 0$$

$$\sum_{a \in \mathcal{E}_-} y_a \leq 0$$

We will use the notation $y_a^+ := \max(y_a, 0)$. Observe that

$$\begin{aligned}\sum_{b\in\mathcal{E}_0} y_b^+ &\geq \sum_{b\in\mathcal{E}_0} y_b \\ &= \sum_{b\in\mathcal{B}} y_b - \sum_{a\in\mathcal{E}} y_a \\ &= 1 - \sum_{a\in\mathcal{E}_+} y_a - \sum_{a\in\mathcal{E}_-} y_a \\ &\geq 1\end{aligned}$$

Define $\xi \in \Delta\mathcal{B}$ by

$$\xi_a := \begin{cases} \frac{y_a^+}{\sum_{b\in\mathcal{E}_0} y_b^+} & \text{if } a \in \mathcal{E}_0 \\ 0 & \text{if } a \in \mathcal{E} \end{cases}$$

We have $\xi \in \mathcal{U}^+$, since both sides of equation (12) vanish when substituing ξ. Hence, we can apply Lemma 36 to conclude

$$\begin{aligned}d_1\left(y, \mathcal{U}^+\right) &\leq \|y - \xi\|_1 \\ &= d_1\left(y, \Delta\mathcal{B}\right)\end{aligned}$$

The second case is when

$$\sum_{a\in\mathcal{E}_+} y_a \geq 0$$

$$\sum_{a\in\mathcal{E}_-} y_a \geq 0$$

$$\sum_{a\in\mathcal{E}_0} y_a \geq 0$$

Define $\xi \in \mathbb{R}^{\mathcal{B}}$ by

$$\xi_a := \begin{cases} \frac{\sum_{b\in\mathcal{E}_+} y_b}{\sum_{b\in\mathcal{E}_+} y_b^+} \cdot y_a^+ & \text{if } a \in \mathcal{E}_+ \\ \\ \frac{\sum_{b\in\mathcal{E}_-} y_b}{\sum_{b\in\mathcal{E}_-} y_b^+} \cdot y_a^+ & \text{if } a \in \mathcal{E}_- \\ \\ \frac{\sum_{b\in\mathcal{E}_0} y_b}{\sum_{b\in\mathcal{E}_0} y_b^+} \cdot y_a^+ & \text{if } a \in \mathcal{E}_0 \end{cases}$$

Here, if the denominator vanishes then y_a^+ must vanish as well, as we interpret the entire expression as 0.

Let $\mathcal{E}_*$ be any of $\mathcal{E}_+$, $\mathcal{E}_-$, $\mathcal{E}_0$. If $\sum_{b\in\mathcal{E}_*} y_b^+ > 0$ then,

$$\begin{aligned}\sum_{a\in\mathcal{E}_*} \xi_a &= \sum_{a\in\mathcal{E}_*} \frac{\sum_{b\in\mathcal{E}_*} y_b}{\sum_{b\in\mathcal{E}_*} y_b^+} \cdot y_a^+ \\ &= \frac{\sum_{b\in\mathcal{E}_*} y_b}{\sum_{b\in\mathcal{E}_*} y_b^+} \cdot \sum_{a\in\mathcal{E}_*} y_a^+ \\ &= \sum_{b\in\mathcal{E}_*} y_b\end{aligned}$$

On the other hand, if $\sum_{b\in\mathcal{E}_*} y_b^+ = 0$ then

$$\begin{aligned}0 &= \sum_{b\in\mathcal{E}_*} y_b^+ \\ &\geq \sum_{b\in\mathcal{E}_*} y_b \\ &\geq 0\end{aligned}$$

Here, we used the case conditions on the last line. We conclude that $\sum_{b\in\mathcal{E}_*} y_b = 0$ and therefore

$$\begin{aligned}\sum_{a\in\mathcal{E}_*} \xi_a &= \sum_{a\in\mathcal{E}_*} 0 \\ &= 0 \\ &= \sum_{b\in\mathcal{E}_*} y_b\end{aligned}$$

We got that *either* way

$$\sum_{a\in\mathcal{E}_*} \xi_a = \sum_{b\in\mathcal{E}_*} y_b$$

It follows that,

$$\begin{aligned}\sum_{a\in\mathcal{B}} \xi_a &= \sum_{a\in\mathcal{E}_+} \xi_a + \sum_{a\in\mathcal{E}_-} \xi_a + \sum_{a\in\mathcal{E}_0} \xi_a \\ &= \sum_{b\in\mathcal{E}_+} y_b + \sum_{b\in\mathcal{E}_-} y_b + \sum_{b\in\mathcal{E}_0} y_b \\ &= \sum_{b\in\mathcal{B}} y_b \\ &= 1\end{aligned}$$

Also, the case conditions imply that $\xi_a \geq 0$. We conclude that $\xi \in \Delta\mathcal{B}$.

Now, let's check equation (12) for ξ:

$$
\begin{aligned}
(1-p)\sum_{a\in\mathcal{E}_+}\xi_a &= (1-p)\sum_{b\in\mathcal{E}_+} y_b \\
&= p\sum_{b\in\mathcal{E}_-} y_b \\
&= p\sum_{a\in\mathcal{E}_-} \xi_a
\end{aligned}
$$

Here, we used equation (12) for y on the second line.

Therefore, $\xi \in \mathcal{U}\cap\Delta\mathcal{B}$. Moreover, it's easy to see that Lemma 36 is applicable and hence,

$$
\begin{aligned}
d_1\left(y,\mathcal{U}^+\right) &\le \|y-\xi\|_1 \\
&= d_1\left(y,\Delta\mathcal{B}\right)
\end{aligned}
$$

The third case is when

$$
\sum_{a\in\mathcal{E}_+} y_a \ge 0
$$

$$
\sum_{a\in\mathcal{E}_-} y_a \ge 0
$$

$$
\sum_{a\in\mathcal{E}_0} y_a < 0
$$

Define $\xi \in \mathbb{R}^{\mathcal{B}}$ by

$$
\xi_a := \begin{cases}
\frac{p}{\sum_{b\in\mathcal{E}_+} y_b^+}\cdot y_a^+ & \text{if } a\in\mathcal{E}_+ \\[2em]
\frac{1-p}{\sum_{b\in\mathcal{E}_-} y_b^+}\cdot y_a^+ & \text{if } a\in\mathcal{E}_- \\[2em]
0 & \text{if } a\in\mathcal{E}_0
\end{cases}
$$

Here, if the denominator vanishes then y_a^+ must vanish as well, as we interpret the entire expression as 0.

If $\sum_{b\in\mathcal{E}_+} y_b^+ > 0$ then,

$$
\begin{aligned}
\sum_{a\in\mathcal{E}_+}\xi_a &= \sum_{a\in\mathcal{E}_+}\frac{p}{\sum_{b\in\mathcal{E}_+} y_b^+}\cdot y_a^+ \\
&= \frac{p}{\sum_{b\in\mathcal{E}_+} y_b^+}\cdot\sum_{a\in\mathcal{E}_+} y_a^+ \\
&= p
\end{aligned}
$$

On the other hand, if $\sum_{b\in\mathcal{E}_+} y_b^+ = 0$ then, like in the second case, $\sum_{a\in\mathcal{E}_+} y_a = 0$. By equation (12) this implies that either $p = 0$ or $\sum_{a\in\mathcal{E}_-} y_a = 0$. However, the latter would imply $\sum_{a\in\mathcal{E}_0} y_a = 1$ which contradicts the case conditions. Therefore, $p = 0$ and in particular

$$\sum_{a\in\mathcal{E}_+} \xi_a = p$$

We got that this identity holds either way. By analogous reasoning,

$$\sum_{a\in\mathcal{E}_-} \xi_a = 1 - p$$

It follows that,

$$\begin{aligned}\sum_{a\in\mathcal{B}} \xi_a &= \sum_{a\in\mathcal{E}_+} \xi_a + \sum_{a\in\mathcal{E}_-} \xi_a + \sum_{a\in\mathcal{E}_0} \xi_a \\ &= p + 1 - p \\ &= 1\end{aligned}$$

We conclude that $\xi \in \Delta\mathcal{B}$.

Now, let's check equation (12) for ξ:

$$\begin{aligned}(1-p)\sum_{a\in\mathcal{E}_+} \xi_a &= (1-p)p \\ &= p\sum_{a\in\mathcal{E}_-} \xi_a\end{aligned}$$

Therefore, $\xi \in \mathcal{U}^+$.

We have,

$$\begin{aligned}\sum_{a\in\mathcal{E}_+} y_a^+ &\geq \sum_{a\in\mathcal{E}_+} y_a \\ &= p\sum_{a\in\mathcal{E}_+} y_a + (1-p)\sum_{a\in\mathcal{E}_+} y_a \\ &= p\sum_{a\in\mathcal{E}_+} y_a + p\sum_{a\in\mathcal{E}_-} y_a \\ &= p\left(\sum_{a\in\mathcal{E}_+} y_a + \sum_{a\in\mathcal{E}_-} y_a\right) \\ &= p\left(1 - \sum_{a\in\mathcal{E}_0} y_a\right) \\ &\geq p\end{aligned}$$

Here, we used equation (12) on the second line and the case conditions on the last line. By analogous reasoning,

$$\sum_{a \in \mathcal{E}_-} y_a^+ \geq 1 - p$$

These inequalities allow us to apply Lemma 36 and conclude that

$$\begin{aligned} d_1\left(y, \mathcal{U}^+\right) &\leq \|y - \xi\|_1 \\ &= d_1\left(y, \Delta\mathcal{B}\right) \end{aligned}$$

■

D.2 System of Absolute Probabilities

The following lemma will be needed for the proof of Proposition 10. It shows that, given a probability distribution ψ over the subsets of a fixed finite set $\mathcal{F}$, the probability $\alpha_i(\psi)$ that a particular element i is in the set can be modified to any other value p without affecting the probabilities of other elements $j \neq i$, in such a manner that the total variation distance between the new distribution ψ' and the original ψ is equal to the probability shift $|\alpha_i(\psi) - p|$.

Lemma 37 *Let $\mathcal{F}$ be a finite set. For every $i \in \mathcal{F}$, define $\alpha_i : \Delta 2^{\mathcal{F}} \to [0, 1]$ by*

$$\alpha_i(\psi) := \Pr_{A \sim \psi}[i \in A]$$

Consider any $i \in \mathcal{F}$, $\psi \in \Delta 2^{\mathcal{F}}$ and $p \in [0, 1]$. Then, there exists $\psi' \in \Delta 2^{\mathcal{F}}$ s.t. the following conditions hold:

- $\alpha_i(\psi') = p$
- *For any $j \in \mathcal{F} \setminus i$, $\alpha_j(\psi') = \alpha_j(\psi)$*
- $\|\psi - \psi'\|_1 = 2\,|\alpha_i(\psi) - p|$

Proof We can assume that $p \geq \alpha_i(\psi)$ without loss of generality: otherwise, we can apply to ψ the bijection $2^{\mathcal{F}} \to 2^{\mathcal{F}}$ defined by taking symmetric difference with $\{i\}$ and change p to $1 - p$. Define $f : 2^{\mathcal{F}} \to 2^{\mathcal{F}}$ by

$$f(A) := A \cup \{i\}$$

Set ψ' to

$$\psi' := \frac{(1 - p)\psi + (p - \alpha_i(\psi))\, f_*\psi}{1 - \alpha_i(\psi)}$$

This is a convex combination of $f_*\psi$ and ψ and hence $\psi' \in \Delta 2^{\mathcal{F}}$. Let's verify the conditions.

For the first condition, we have

$$\begin{aligned}
\alpha_i\left(\psi'\right) &= \frac{(1-p)\alpha_i(\psi) + \left(p - \alpha_i(\psi)\right)\alpha_i(f_*\psi)}{1-\alpha_i(\psi)} \\
&= \frac{\alpha_i(\psi) - p\alpha_i(\psi) + p - \alpha_i(\psi)}{1-\alpha_i(\psi)} \\
&= \frac{p\left(1-\alpha_i(\psi)\right)}{1-\alpha_i(\psi)} \\
&= p
\end{aligned}$$

Here, we used that $i \in f(A)$ for any $A \in 2^{\mathcal{F}}$ to conclude that $\alpha_i(f_*\psi) = 1$.

For the second condition,

$$\begin{aligned}
\alpha_j\left(\psi'\right) &= \frac{(1-p)\alpha_j(\psi) + \left(p - \alpha_i(\psi)\right)\alpha_j(f_*\psi)}{1-\alpha_i(\psi)} \\
&= \frac{(1-p)\alpha_j(\psi) + \left(p - \alpha_i(\psi)\right)\alpha_j(\psi)}{1-\alpha_i(\psi)} \\
&= \frac{\left(1-p+p-\alpha_i(\psi)\right)\alpha_j(\psi)}{1-\alpha_i(\psi)} \\
&= \frac{\left(1-\alpha_i(\psi)\right)\alpha_j(\psi)}{1-\alpha_i(\psi)} \\
&= \alpha_j(\psi)
\end{aligned}$$

Here, we used that $j \in f(A)$ if and only if $j \in A$ to conclude that $\alpha_j(f_*\psi) = \alpha_j(\psi)$.

For the third condition, first observe that

$$
\begin{aligned}
\|\psi - f_*\psi\|_1 &= \sum_{A\subseteq\mathcal{F}} |\psi_A - f_*\psi_A| \\
&= \sum_{A\subseteq\mathcal{F}} \left|\psi_A - \sum_{B\subseteq\mathcal{F}:f(B)=A} \psi_B\right| \\
&= \sum_{A\subseteq\mathcal{F}:i\in A} \left|\psi_A - \sum_{B\subseteq\mathcal{F}:f(B)=A} \psi_B\right| + \\
&\quad \sum_{A\subseteq\mathcal{F}:i\notin A} \left|\psi_A - \sum_{B\subseteq\mathcal{F}:f(B)=A} \psi_B\right| \\
&= \sum_{A\subseteq\mathcal{F}:i\in A} \left|\psi_A - \left(\psi_A + \psi_{A\setminus i}\right)\right| + \sum_{A\subseteq\mathcal{F}:i\notin A} \psi_A \\
&= \sum_{A\subseteq\mathcal{F}:i\in A} \psi_{A\setminus i} + \sum_{A\subseteq\mathcal{F}:i\notin A} \psi_A \\
&= 2\sum_{A\subseteq\mathcal{F}:i\notin A} \psi_A \\
&= 2\left(1-\alpha_i(\psi)\right)
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\left\|\psi - \psi'\right\|_1 &= \left\|\psi - \frac{(1-p)\psi + (p-\alpha_i(\psi))\, f_*\psi}{1-\alpha_i(\psi)}\right\|_1 \\
&= \left\|\frac{(1-p)\psi + (p-\alpha_i(\psi))\,\psi}{1-\alpha_i(\psi)} - \frac{(1-p)\psi + (p-\alpha_i(\psi))\, f_*\psi}{1-\alpha_i(\psi)}\right\|_1 \\
&= \left\|\frac{(1-p)(\psi-\psi) + (p-\alpha_i(\psi))\,(\psi - f_*\psi)}{1-\alpha_i(\psi)}\right\|_1 \\
&= \frac{p-\alpha_i(\psi)}{1-\alpha_i(\psi)} \cdot \|\psi - f_*\psi\|_1 \\
&= \frac{p-\alpha_i(\psi)}{1-\alpha_i(\psi)} \cdot 2\left(1-\alpha_i(\psi)\right) \\
&= 2\left(p-\alpha_i(\psi)\right)
\end{aligned}
$$

■

Another lemma we need for Proposition 10: given two probability distributions on a finite set B, if we somehow "lift" one of them through a surjection, then we can also lift the other s.t. the total variation distance between them is preserved. This is achieved by extracting conditional distributions on the fibers from the lift, and combining them with the other distribution on B.

Lemma 38 *Let A, B be finite sets, $f : A \to B$ a surjection, $\xi \in \Delta A$ and $\psi \in \Delta B$. Then, there exists $\xi' \in \Delta A$ s.t. $f_*\xi' = \psi$ and*

$$\left\|\xi - \xi'\right\|_1 = \|f_*\xi - \psi\|_1$$

Proof Set $\xi' \in \mathbb{R}^A$ to

$$\xi'_i := \frac{\psi_{f(i)}\xi_i}{\sum_{j \in f^{-1}(f(i))} \xi_j}$$

Let's verify that $\xi' \in \Delta A$:

$$\begin{aligned}\sum_{i \in \mathcal{A}} \xi_i &= \sum_{i \in \mathcal{A}} \frac{\psi_{f(i)}\xi_i}{\sum_{j \in f^{-1}(f(i))} \xi_j} \\ &= \sum_{k \in B} \sum_{i \in f^{-1}(k)} \frac{\psi_k \xi_i}{\sum_{j \in f^{-1}(k)} \xi_j} \\ &= \sum_{k \in B} \frac{\psi_k}{\sum_{j \in f^{-1}(k)} \xi_j} \cdot \sum_{i \in f^{-1}(k)} \xi_i \\ &= \sum_{k \in B} \psi_k \\ &= 1\end{aligned}$$

Let's verify that $f_*\xi' = \psi$:

$$\begin{aligned}f_*\xi'_k &= \sum_{i \in f^{-1}(k)} \xi'_i \\ &= \sum_{i \in f^{-1}(k)} \frac{\psi_k \xi_i}{\sum_{j \in f^{-1}(k)} \xi_j} \\ &= \frac{\psi_k}{\sum_{j \in f^{-1}(k)} \xi_j} \cdot \sum_{i \in f^{-1}(k)} \xi_i \\ &= \psi_k\end{aligned}$$

Finally, we have

$$\begin{aligned}
\left\|\xi-\xi'\right\|_1 &= \sum_{i\in A}\left|\xi_i-\xi'_i\right| \\
&= \sum_{i\in A}\left|\xi_i-\frac{\psi_{f(i)}\xi_i}{\sum_{j\in f^{-1}(f(i))}\xi_j}\right| \\
&= \sum_{i\in A}\xi_i\left|1-\frac{\psi_{f(i)}}{\sum_{j\in f^{-1}(f(i))}\xi_j}\right| \\
&= \sum_{k\in B}\sum_{i\in f^{-1}(k)}\xi_i\left|1-\frac{\psi_k}{\sum_{j\in f^{-1}(k)}\xi_j}\right| \\
&= \sum_{k\in B}\sum_{i\in f^{-1}(k)}\xi_i\cdot\frac{1}{\sum_{j\in f^{-1}(k)}\xi_j}\cdot\left|\sum_{j\in f^{-1}(k)}\xi_j-\psi_k\right| \\
&= \sum_{k\in B}\frac{1}{\sum_{j\in f^{-1}(k)}\xi_j}\cdot\left|\sum_{j\in f^{-1}(k)}\xi_j-\psi_k\right|\sum_{i\in f^{-1}(k)}\xi_i \\
&= \sum_{k\in B}\left|\sum_{j\in f^{-1}(k)}\xi_j-\psi_k\right| \\
&= \left\|f_*\xi-\psi\right\|_1
\end{aligned}$$

■

We are now ready to prove Proposition 10. The idea is as follows. Starting from some $y \in \mathcal{U}^\flat \setminus \Delta\mathcal{B}$, we first use Proposition 11 to "project" it to $\Delta\mathcal{B}$ while preserving *one* of the probabilities in the system and possibly "spoiling" the others. We then "fix" each of the other probabilities sequentially using the two previous lemmas, until we finally get some $\xi \in \mathcal{U}^+$ whose distance from y can be bounded by the distances incurred at each step of this process.

Proof [Proof of Proposition 10]Assume without loss of generality that $\mathcal{F} = [n]$ for some $n \geq 1$ and consider any $y \in \mathcal{U}^\flat \setminus \Delta\mathcal{B}$. Let α be defined as in Lemma 37, denote $\delta := d_1(y, \Delta\mathcal{B})$ and define $\mathcal{U}_0$ by

$$\mathcal{U}_0 := \left\{y \in \mathbb{R}^{\mathcal{B}} \,\middle|\, \alpha_0\left(f_*y\right) = p_0\right\}$$

Notice that $\mathcal{U} \subseteq \mathcal{U}_0$ and in particular $y \in \mathcal{U}_0$. By Proposition 11, $\sin(\mathcal{U}_0^\flat, \Delta\mathcal{B}) = 1$. Hence, there exists $\xi_1 \in \Delta\mathcal{B}$ s.t.

$$\alpha_0\left(f_*\xi_1\right) = p_0$$

$$\left\|y-\xi_1\right\|_1 = \delta$$

Denote $\psi_1 := f_*\xi_1$. Observe that for any $0 < i < n$,

$$\begin{aligned}
|\alpha_i(\psi_1) - p_i| &= |\alpha_i(\psi_1) - \alpha_i(f_* y)| \\
&\leq \frac{1}{2}\|\psi_1 - f_* y\|_1 \\
&= \frac{1}{2}\|f_*\xi_1 - f_* y\|_1 \\
&\leq \frac{1}{2}\|\xi_1 - y\|_1 \\
&= \frac{1}{2}\delta
\end{aligned}$$

Here, we extended the definition of α_i to $\mathbb{R}^{2^{\mathcal{F}}}$ such as to make it a linear functional. Applying Lemma 37 by induction, we construct $\{\psi_i \in \Delta 2^{\mathcal{F}}\}_{2 \leq i \leq n}$ s.t.:

- For any $1 \leq i \leq n$ and $0 \leq j < i$, $\alpha_j(\psi_i) = p_j$.
- For any $1 \leq i \leq n$ and $i \leq j < n$, $\alpha_j(\psi_i) = \alpha_j(\psi_1)$ and in particular $|\alpha_j(\psi_i) - p_j| \leq \frac{1}{2}\delta$.
- For any $1 \leq i < n$, $\|\psi_i - \psi_{i+1}\|_1 \leq \delta$.

By the triangle inequality,

$$\|\psi_1 - \psi_n\|_1 \leq (n-1)\delta$$

Applying Lemma 38 we get $\xi \in \Delta\mathcal{B}$ s.t.

- $f_*\xi = \psi_n$
- $\|\xi_1 - \xi\| \leq (n-1)\delta$

By the construction of ψ_n, the first item above implies that $\xi \in \mathcal{U}$. We get that

$$\begin{aligned}
d_1\left(y, \mathcal{U}^+\right) &\leq \|y - \xi\|_1 \\
&\leq \|y - \xi_1\|_1 + \|\xi_1 - \xi\|_1 \\
&= \delta + (n-1)\delta \\
&= n\delta \\
&= |\mathcal{F}|\, d_1(y, \Delta\mathcal{B})
\end{aligned}$$

By Lemma 33, this implies

$$\sin\left(\mathcal{U}^\flat, \Delta\mathcal{B}\right) \geq \frac{1}{|\mathcal{F}|}$$

■

D.3 S for Chain of Conditional Probabilities

Now, we start working towards the proof of Proposition 12. The following shows that if some $a^* \in \mathcal{B}$ is in the "support" of some $\mathcal{U} \in \mathrm{Gr}^+\left(\mathbb{R}^{\mathcal{B}}\right)$, then it is also in the "support" of $\mathcal{U}^\flat$.

Lemma 39 *Let $\mathcal{B}$ be a finite set, $\mathcal{U} \subseteq \mathbb{R}^{\mathcal{B}}$ a linear subspace. Assume there is some $y \in \mathcal{U}$ s.t. $\sum_a y_a = 1$. Let $\mathcal{E} \subseteq \mathcal{B}$ be the minimal set s.t. $\mathcal{U} \subseteq \mathbb{R}^{\mathcal{E}}$. Then, for any $a^* \in \mathcal{E}$, there exists $y^* \in \mathcal{U}$ s.t. $\sum_a y^*_a = 1$ and $y^*_{a^*} \neq 0$.*

Proof Let $y \in \mathcal{U}$ be s.t. $\sum_a y_a = 1$. In case $y_{a^*} \neq 0$, take $y^* := y$. In case $y_{a^*} = 0$, let $y' \in \mathcal{U}$ be s.t. $y'_{a^*} \neq 0$ (it exists because $a^* \in \mathcal{E}$). In case $\sum_a y'_a \neq 0$, take

$$y^* := \frac{y'}{\sum_a y'_a}$$

In case $\sum_a y'_a = 0$, take $y^* := y + y'$. We get

$$\begin{aligned}\sum_a y^*_a &= \sum_a y_a + \sum_a y'_a \\ &= 1 + 0 \\ &= 1\end{aligned}$$

and,

$$\begin{aligned}y^*_{a^*} &= y_{a^*} + y'_{a^*} \\ &= y'_{a^*} \\ &\neq 0\end{aligned}$$

■

When analyzing sines related to Proposition 12, it will be inconvenient to consider completely arbitrary $y \in \mathcal{U}^\flat \setminus \Delta\mathcal{B}$. Instead, we will restrict y to a certain dense subset of "non-degenerate" vectors.

Lemma 40 *Let $\mathcal{G}_0$, $\mathcal{G}_1$ be finite sets, $\mathcal{B} := \mathcal{G}_0 \times \mathcal{G}_1$, $\mathcal{U}_0 \in \mathrm{Gr}^+\left(\mathbb{R}^{\mathcal{G}_0}\right)$ and $\mathcal{U}_1 : \mathcal{G}_0 \to \mathrm{Gr}^+\left(\mathbb{R}^{\mathcal{G}_1}\right)$. Let $\mathcal{E} \subseteq \mathcal{G}_0$ be the minimal set s.t. $\mathcal{U}_0 \subseteq \mathbb{R}^{\mathcal{E}}$. Given $y \in \mathbb{R}^{\mathcal{B}}$, define $y^0 \in \mathbb{R}^{\mathcal{G}_0}$ by*

$$y^0_a := \sum_{b \in \mathcal{G}_1} y_{ab}$$

Given also $a \in \mathcal{G}_0$, define $y^a \in \mathbb{R}^{\mathcal{G}_1}$ by

$$y^a_b := y_{ab}$$

Define $\mathcal{U}$ by

$$\mathcal{U} := \left\{ y \in \mathbb{R}^{\mathcal{B}} \,\middle|\, y^0 \in \mathcal{U}_0,\ \forall a \in \mathcal{E} : y^a \in \mathcal{U}_1(a) \text{ and } \forall a \in \mathcal{G}_0 \setminus \mathcal{E} : y^a = 0 \right\}$$

Define also $\mathcal{U}^*$ by

$$\mathcal{U}^* := \left\{ y \in \mathcal{U}^\flat \;\middle|\; \exists v \in \prod_{a \in \mathcal{G}_0} \mathcal{U}_1(a)^\flat \; \forall a \in \mathcal{G}_0 : y^a = y^0_a v(a) \right\}$$

Then, $\mathcal{U}^*$ is dense in $\mathcal{U}^\flat$.

Proof Consider any $a_0 \in \mathcal{E}$. By Lemma 39, there is some $\hat{u} \in \mathcal{U}_0^\flat$ s.t. $\hat{u}_{a_0} \neq 0$. Choose any $\hat{v} \in \prod_{a \in \mathcal{G}_0} \mathcal{U}_1(a)^\flat$ and define $\hat{y} \in \mathbb{R}^\mathcal{B}$ by $\hat{y}_{ab} := \hat{u}_a \hat{v}(a)_b$. We have

$$\begin{aligned}
\hat{y}^0_a &= \sum_{b \in \mathcal{G}_1} \hat{y}_{ab} \\
&= \sum_{b \in \mathcal{G}_1} \hat{u}_a \hat{v}(a)_b \\
&= \hat{u}_a \sum_{b \in \mathcal{G}_1} \hat{v}(a)_b \\
&= \hat{u}_a
\end{aligned}$$

Hence, $\hat{y}^0 = \hat{u} \in \mathcal{U}_0$. Moreover, for any $a \in \mathcal{E}$, $\hat{y}^a = \hat{u}_a \hat{v}(a) \in \mathcal{U}_1(a)$. And, for any $a \in \mathcal{G}_0 \setminus \mathcal{E}$, $\hat{u}_a = 0$ and hence $\hat{y}^a = 0$. We conclude that $\hat{y} \in \mathcal{U}$. Also,

$$\begin{aligned}
\sum_{a,b} \hat{y}_{ab} &= \sum_{a,b} \hat{u}_a \hat{v}(a)_b \\
&= \sum_a \hat{u}_a \sum_b \hat{v}(a)_b \\
&= \sum_a \hat{u}_a \\
&= 1
\end{aligned}$$

Therefore, $\hat{y} \in \mathcal{U}^\flat$. And, $\hat{y}^0_{a_0} = \hat{u}_{a_0} \neq 0$. Since $y^0_{a_0} = 0$ is a linear condition on y, it follows $y^0_{a_0} \neq 0$ for all $y \in \mathcal{U}^\flat$ except a set of measure zero w.r.t. the intrinsic Lebesgue measure of $\mathcal{U}^\flat$. Hence, for almost all (in the same sense) $y \in \mathcal{U}^\flat$ it holds that for any $a \in \mathcal{E}$, $y^0_a \neq 0$.

It remains to show that any y with the latter property is in $\mathcal{U}^*$. Indeed, consider some such $y \in \mathcal{U}^\flat$. Choose some $v' \in \prod_a \mathcal{U}_1(a)^\flat$ and define $v \in \prod_a \mathbb{R}^{\mathcal{G}_1}$ by

$$v(a) := \begin{cases} \frac{y^a}{y^0_a} & \text{if } a \in \mathcal{E} \\ v'(a) & \text{if } a \notin \mathcal{E} \end{cases}$$

Since for any $a \in \mathcal{G}_1$, $y^a \in \mathcal{U}_1(a)$, it follows that $v \in \prod_a \mathcal{U}_1(a)^\flat$. Finally, let's check that $y^a = y^0_a v(a)$. For $a \in \mathcal{E}$ we have

$$\begin{aligned}
y^0_a v(a) &= y^0_a \cdot \frac{y^a}{y^0_a} \\
&= y^a
\end{aligned}$$

For $a \notin \mathcal{E}$, $y^0_a = 0$ since $y^0 \in \mathcal{U}_0$. And, $y^a = 0$ since $y \in \mathcal{U}$. Hence $y^0_a v(a) = 0 = y^a$. ■

The total variation distance between two probability distributions on a product space can be bounded in terms of the total variation distances between the associated marginal and conditional distributions. The following is an extension of this bound to the affine hull of the simplex of distributions.

Lemma 41 *Let $\mathcal{G}_0$, $\mathcal{G}_1$ be finite sets, $u, u' \in \mathbb{R}^{\mathcal{G}_0}$, $v : \mathcal{G}_0 \to \mathbb{R}^{\mathcal{G}_1}$, $v' : \mathcal{G}_0 \to \Delta\mathcal{G}_1$. Assume that $\sum_a u_a = 1$, $\sum_a u'_a = 1$ and $\sum_b v(a)_b = 1$. Define $y, y' \in \mathbb{R}^{\mathcal{G}_0 \times \mathcal{G}_1}$ by $y_{ab} := u_a v(a)_b$ and $y'_{ab} := u'_a v'(a)_b$. Then,*

$$\left\|y - y'\right\|_1 \leq \left\|u - u'\right\|_1 + \sum_{a \in \mathcal{G}_0} |u_a| \cdot \left\|v(a) - v(a)'\right\|_1$$

Proof We have,

$$\begin{aligned}
\left\|y - y'\right\|_1 &= \sum_{a,b} \left|y_{ab} - y'_{ab}\right| \\
&= \sum_{a,b} \left|u_a v(a)_b - u'_a v'(a)_b\right| \\
&= \sum_{a,b} \left|u_a v(a)_b - u_a v'(a)_b + u_a v'(a)_b - u'_a v'(a)_b\right| \\
&= \sum_{a,b} \left|u_a \left(v(a)_b - v'(a)_b\right) + \left(u_a - u'_a\right) v'(a)_b\right| \\
&\leq \sum_{a,b} \left|u_a \left(v(a)_b - v'(a)_b\right)\right| + \sum_{a,b} \left|\left(u_a - u'_a\right) v'(a)_b\right| \\
&= \sum_a |u_a| \sum_b \left|v(a)_b - v'(a)_b\right| + \sum_a \left|u_a - u'_a\right| \sum_b v'(a)_b \\
&= \sum_a |u_a| \cdot \left\|v(a) - v(a)'\right\| + \left\|u - u'\right\|_1
\end{aligned}$$

■

The following is essentially the special case of Proposition 12 for $n = 2$. The proof proceeds by decomposing $y \in \mathcal{U}^\flat \setminus \Delta\mathcal{B}$ into "marginal" and "conditionals" (using the non-degeneracy condition afforded by Lemma 40) and then by "projecting" each of them to the intersection of its respective subspace with its respective simplex and combining the resulting distributions.

Lemma 42 *In the setting of Lemma 40,*

$$\sin\left(\mathcal{U}^\flat, \Delta\mathcal{B}\right) \geq \min\left(\sin\left(\mathcal{U}_0^\flat, \Delta\mathcal{G}_0\right), \min_{a \in \mathcal{G}_0} \sin\left(\mathcal{U}_1(a)^\flat, \Delta\mathcal{G}_1\right)\right)$$

Proof Consider any $y \in \mathcal{U}^*$. Let $v \in \prod_a \mathcal{U}_1(a)^\flat$ be s.t. $y^a = y^0_a v(a)$. By Lemma 33, the relevant norms for our purpose are ℓ_1. By Lemma 34,

$$\begin{aligned}
d_1(y, \Delta\mathcal{B}) &= \sum_{a,b} |y_{ab}| - 1 \\
&= \sum_{a,b} \left|y^0_a v(a)_b\right| - 1 \\
&= \sum_a \left|y^0_a\right| \sum_b |v(a)_b| - \sum_a \left|y^0_a\right| + \sum_a \left|y^0_a\right| - 1 \\
&= \sum_a \left|y^0_a\right| \left(\sum_b |v(a)_b| - 1\right) + \sum_a \left|y^0_a\right| - 1 \\
&= \sum_a \left|y^0_a\right| d_1(v(a), \Delta\mathcal{G}_1) + d_1\left(y^0, \Delta\mathcal{G}_0\right)
\end{aligned}$$

Let $\xi^0 \in \Delta\mathcal{G}_0$ and $\xi^1 : \mathcal{G}_0 \to \Delta\mathcal{G}_1$ be

$$\xi^0 := \underset{\xi \in \mathcal{U}_0 \cap \Delta\mathcal{G}_0}{\operatorname{argmin}} \left\|y^0 - \xi\right\|_1$$

$$\xi^1(a) := \underset{\xi \in \mathcal{U}_1(a) \cap \Delta\mathcal{G}_1}{\operatorname{argmin}} \|v(a) - \xi\|_1$$

Define $\xi^* \in \mathcal{U} \cap \Delta\mathcal{B}$ by $\xi^*_{ab} := \xi^0_a \xi^1(a)_b$. Denote

$$S_{\min} := \min\left(\sin\left(\mathcal{U}^\flat_0, \Delta\mathcal{G}_0\right), \min_a \sin\left(\mathcal{U}_1(a)^\flat, \Delta\mathcal{G}_1\right)\right)$$

By Lemma 41,

$$\begin{aligned}
d_1(y, \mathcal{U} \cap \Delta\mathcal{B}) &\leq \|y - \xi^*\| \\
&\leq \sum_a \left|y^0_a\right| \cdot \left\|v(a) - \xi^1(a)\right\|_1 + \left\|y^0 - \xi^0\right\|_1 \\
&= \sum_a \left|y^0_a\right| d_1(v(a), \mathcal{U}_1(a) \cap \Delta\mathcal{G}_1) + d_1\left(y^0, \mathcal{U}_0 \cap \Delta\mathcal{G}_0\right) \\
&\leq \sum_a \left|y^0_a\right| \cdot \frac{d_1(v(a), \Delta\mathcal{G}_1)}{\sin\left(\mathcal{U}_1(a)^\flat, \Delta\mathcal{G}_1\right)} + \frac{d_1\left(y^0, \Delta\mathcal{G}_0\right)}{\sin\left(\mathcal{U}^\flat_0, \Delta\mathcal{G}_0\right)} \\
&\leq \sum_a \left|y^0_a\right| \cdot \frac{d_1(v(a), \Delta\mathcal{G}_1)}{S_{\min}} + \frac{d_1\left(y^0, \Delta\mathcal{G}_0\right)}{S_{\min}} \\
&= \frac{1}{S_{\min}} \left(\sum_a \left|y^0_a\right| d_1(v(a), \Delta\mathcal{G}_1) + d_1\left(y^0, \Delta\mathcal{G}_0\right)\right) \\
&= \frac{1}{S_{\min}} \cdot d_1(y, \Delta\mathcal{B})
\end{aligned}$$

This holds for any $y \in \mathcal{U}^*$. And, by Lemma 40, such y are dense in $\mathcal{U}^\flat$ and hence

$$\sin\left(\mathcal{U}^\flat, \Delta\mathcal{B}\right) \geq S_{\min}$$

■

In the following, we will use the following shorthand notation for portions of tuples:

$$x_{:i} := x_0 x_1 \ldots x_{i-1}$$

$$x_{i:j} := x_i x_{i+1} \ldots x_{j-1}$$

We will require a characterization of the "support" of $\mathcal{U}$ from Proposition 12 (we state it only in the direction we need but the converse also holds).

Lemma 43 *In the setting of Proposition 12, let $\mathcal{E}$ be the minimal subset of $\mathcal{B}$ s.t. $\mathcal{U} \subseteq \mathbb{R}^{\mathcal{B}}$. Let $a^* \in \mathcal{B} \setminus \mathcal{E}$. Then, there exists an integer $i < n$ s.t.*

$$a^*_i \notin \mathcal{E}_i\left(a^*_{:i}\right)$$

Proof Assume to the contrary that for any $i < n$,

$$a^*_i \in \mathcal{E}_i\left(a^*_{:i}\right)$$

For every $i < n$, choose some $u^i : \mathcal{G}^i_- \to \mathbb{R}^{\mathcal{G}_i}$ s.t.

- For any $a \in \mathcal{G}^i_-$, $u^i(a) \in \mathcal{U}^\flat_i$.
- $u^i(a^*_{:i})_{a^*_i} \neq 0$ (possible by Lemma 39).

Define $y \in \mathbb{R}^{\mathcal{B}}$ by

$$y_a := \prod_{i<n} u^i\left(a_{:i}\right)_{a_i}$$

Let's check that $y \in \mathcal{U}$.

For any $i < n$ and $a \in \mathcal{G}^i_-$, we have

$$
\begin{aligned}
y^a_b = \quad & \sum_{c \in \mathcal{G}^i_+} y_{abc} \\
= \quad & \sum_{c \in \mathcal{G}^i_+} \left(\prod_{j<i} u^j\left(a_{:j}\right)_{a_j} \cdot u^i\left(a\right)_b \cdot \prod_{i<j<n} u^j\left(abc_{i+1:j}\right)_{c_j} \right) \\
= \quad & \prod_{j<i} u^j\left(a_{:j}\right)_{a_j} \cdot u^i\left(a\right)_b \cdot \\
& \sum_{c_{i+1}} u^{i+1}\left(ab\right)_{c_{i+1}} \sum_{c_{i+2}} u^{i+2}\left(abc_{i+1}\right)_{c_{i+2}} \ldots \sum_{c_{n-1}} u^{n-1}\left(abc_{i+1:n-1}\right)_{c_{n-1}} \\
= \quad & \prod_{j<i} u^j\left(a_{:j}\right)_{a_j} \cdot u^i\left(a\right)_b \cdot \\
& \sum_{c_{i+1}} u^{i+1}\left(ab\right)_{c_{i+1}} \sum_{c_{i+2}} u^{i+2}\left(abc_{i+1}\right)_{c_{i+2}} \ldots \sum_{c_{n-2}} u^{n-2}\left(abc_{i+1:n-2}\right)_{c_{n-2}} \\
= \quad & \ldots \\
= \quad & \prod_{j<i} u^j\left(a_{:j}\right)_{a_j} \cdot u^i\left(a\right)_b
\end{aligned}
$$

Here, we repeatedly used the fact that $\sum_{c_j} u^j(x)_{c_j} = 1$ to eliminate the nested sums.

The first factor on the right hand side doesn't depend on b, hence y^a is a scalar times $u^i(a)$ and is therefore in $\mathcal{U}_i(a)$. Moreover, for any $b \in \mathcal{G}_i \setminus \mathcal{E}_i\left(a\right)$ and $c \in \mathcal{G}^i_+$, we have

$$
\begin{aligned}
y_{abc} &= \sum_{c \in \mathcal{G}^i_+} \left(\prod_{j<i} u^j\left(a_{:j}\right)_{a_j} \cdot u^i\left(a\right)_b \cdot \prod_{i<j<n} u^j\left(abc_{i+1:j}\right)_{c_j} \right) \\
&= \sum_{c \in \mathcal{G}^i_+} \left(\prod_{j<i} u^j\left(a_{:j}\right)_{a_j} \cdot 0 \cdot \prod_{i<j<n} u^j\left(abc_{i+1:j}\right)_{c_j} \right) \\
&= 0
\end{aligned}
$$

Now, observe that $y_{a^*} \neq 0$ since it is a product of non-vanishing factors. But, this implies $a^* \in \mathcal{E}$, which is a contradiction. ■

We are now ready to prove Proposition 12. The proof works by observing that $\mathcal{U}$ can be described as the recursive application of the construction in Lemma 40 and using Lemma 42 by induction.

Proof [Proof of Proposition 12]We use induction on n.

For $n = 1$, $y \in \mathcal{U}$ if and only if: $y \in \mathcal{U}_0$ and for all $b \in \mathcal{G}_0 \setminus \mathcal{E}_0$, $y_b = 0$. But, the latter condition follows from the former, by definition of $\mathcal{E}_0$. Hence, $\mathcal{U} = \mathcal{U}_0$ and the claim is true.

Now, assume the claim for some $n \geq 1$, and let's show it for $n+1$. For any $i < n$, define

$$
\mathcal{G}^i_{(+)} := \prod_{i<j<n} \mathcal{G}_j
$$

Notice that this is not the same thing as $\mathcal{G}^i_+$, since in this context $\mathcal{G}^i_+ = \prod_{i<j<n+1} \mathcal{G}_j$ (because we are proving the claim for $n+1$). Define also

$$\tilde{\mathcal{G}}_0 := \mathcal{G}^n_-$$

$$\tilde{\mathcal{G}}_1 := \mathcal{G}_n$$

$$\tilde{\mathcal{U}}_0 := \{y \in \mathbb{R}^{\mathcal{G}^n_-} \mid \forall i < n, a \in \mathcal{G}^i_- : y^a \in \mathcal{U}_i(a) \text{ and } \forall b \in \mathcal{G}_i \setminus \mathcal{E}_i\left(a\right), c \in \mathcal{G}^i_{(+)} : y_{abc} = 0\}$$

$$\tilde{\mathcal{U}}_1 := \mathcal{U}_n$$

Given $y \in \mathbb{R}^{\mathcal{B}}$, define $y^0 \in \mathbb{R}^{\mathcal{G}^n_-}$ by

$$y^0_a := \sum_{b \in \mathcal{G}_n} y_{ab}$$

Let $\tilde{\mathcal{E}} \subseteq \mathcal{G}^n_-$ be the minimal set s.t. $\tilde{\mathcal{U}}_0 \subseteq \mathbb{R}^{\tilde{\mathcal{E}}}$. Finally, define

$$\tilde{\mathcal{U}} := \left\{ y \in \mathbb{R}^{\mathcal{B}} \,\middle|\, y^0 \in \tilde{\mathcal{U}}_0, \forall a \in \tilde{\mathcal{E}} : y^a \in \mathcal{U}_n(a) \text{ and } \forall a \in \mathcal{G}^n_- \setminus \tilde{\mathcal{E}} : y^a = 0 \right\}$$

Observe that for any $y \in \mathbb{R}^{\mathcal{B}}$, $i < n$ and $a \in \mathcal{G}^i_-$, we have $y^a = (y^0)^a$ because

$$\begin{aligned} y^a_b &= \sum_{c \in \mathcal{G}^i_+} y_{abc} \\ &= \sum_{d \in \mathcal{G}^i_{(+)}} \sum_{e \in \mathcal{G}_n} y_{abde} \\ &= \sum_{d \in \mathcal{G}^i_{(+)}} y^0_{abd} \\ &= \left(y^0\right)^a_b \end{aligned}$$

Let's show that $\tilde{\mathcal{U}} \subseteq \mathcal{U}$. Consider any $y \in \tilde{\mathcal{U}}$ and let's check that $y \in \mathcal{U}$.

For any $i < n$ and $a \in \mathcal{G}^i_-$, $y^a = (y^0)^a$. Since $y^0 \in \tilde{\mathcal{U}}_0$, we get $(y^0)^a \in \mathcal{U}_i(a)$ and hence $y^a \in \mathcal{U}_i(a)$.

Moreover, for any $b \in \mathcal{G}_i \setminus \mathcal{E}_i\left(a\right)$, $d \in \mathcal{G}^i_{(+)}$ and $y' \in \tilde{\mathcal{U}}_0$, we have $y'_{abd} = 0$. Therefore, $abd \notin \tilde{\mathcal{E}}$. Hence, $y^{abd} = 0$, meaning that for any $e \in \mathcal{G}_n$, $y_{abde} = 0$.

For any $a \in \mathcal{G}^n_-$, if $a \in \tilde{\mathcal{E}}$ then $y^a \in \mathcal{U}_n(a)$. If $a \notin \tilde{\mathcal{E}}$, then $y^a = 0 \in \mathcal{U}_n(a)$. Hence, in either case $y^a \in \mathcal{U}_n(a)$.

Moreover, for any $b \in \mathcal{G}_n \setminus \mathcal{E}_n\left(a\right)$, $y_{ab} = y^a_b = 0$, because $y^a \in \mathcal{U}_n(a)$.

Now, let's show that $\mathcal{U} \subseteq \tilde{\mathcal{U}}$. Consider any $y \in \mathcal{U}$ and let's check that $y \in \tilde{\mathcal{U}}$.

First, we need to check that $y^0 \in \tilde{\mathcal{U}}_0$. For any $i < n$ and $a \in \mathcal{G}_-^i$, $(y^0)^a = y^a \in \mathcal{U}_i(a)$. For any $b \in \mathcal{G}_i \setminus \mathcal{E}_i(a)$ and $d \in \mathcal{G}_{(+)}^i$,

$$\begin{aligned} y^0_{abd} &= \sum_{e \in \mathcal{G}_n} y_{abde} \\ &= \sum_{e \in \mathcal{G}_n} 0 \\ &= 0 \end{aligned}$$

For any $a \in \tilde{\mathcal{E}}$, we have $y^a \in \mathcal{U}_n(a)$ since $y \in \mathcal{U}$. For any $a \in \mathcal{G}_-^n \setminus \tilde{\mathcal{E}}$, Lemma 43 implies that there is some $i < n$ s.t. $a_i \notin \mathcal{E}_i(a_{:i})$. Hence, for any $b \in \mathcal{G}_n$,

$$\begin{aligned} y^a_b &= y_{ab} \\ &= y_{a_{:i} a_i a_{i+1:n} b} \\ &= 0 \end{aligned}$$

We got that $\mathcal{U} = \tilde{\mathcal{U}}$. Hence, Lemma 42 applied to $\tilde{\mathcal{G}}_0, \tilde{\mathcal{G}}_1, \tilde{\mathcal{U}}_0, \tilde{\mathcal{U}}_1$ implies that

$$\sin\left(\mathcal{U}^\flat, \Delta\mathcal{B}\right) \geq \min\left(\sin\left(\tilde{\mathcal{U}}_0^\flat, \Delta\mathcal{G}_-^n\right), \min_{a \in \mathcal{G}_-^n} \sin\left(\mathcal{U}_n(a)^\flat, \Delta\mathcal{G}_n\right)\right)$$

Applying the induction hypothesis to $\tilde{\mathcal{U}}_0$, we conclude

$$\begin{aligned} \sin\left(\mathcal{U}^\flat, \Delta\mathcal{B}\right) &\geq \min\left(\min_{\substack{i<n \\ a \in \mathcal{G}_-^i}} \sin\left(\mathcal{U}_i(a)^\flat, \Delta\mathcal{G}_i\right), \min_{a \in \mathcal{G}_-^n} \sin\left(\mathcal{U}_n(a)^\flat, \Delta\mathcal{G}_n\right)\right) \\ &= \min_{\substack{i<n+1 \\ a \in \mathcal{G}_-^i}} \sin\left(\mathcal{U}_i(a)^\flat, \Delta\mathcal{G}_i\right) \end{aligned}$$

■

D.4 R for Chain of Conditional Probabilities

Now, we move towards the proof of Proposition 13. We will need a technical lemma about the ℓ_1 distances of affine subspaces from the origin.

Lemma 44 *Let $\mathcal{B}$ be a finite set, $\mathcal{U} \subseteq \mathbb{R}^{\mathcal{B}}$ a linear subspace s.t. $\mathcal{U} \cap \Delta\mathcal{B} \neq \varnothing$, $u_0 \in \mathbb{R}^{\mathcal{B}}$ and $t \in \mathbb{R}$. Define $\mu : \mathbb{R}^{\mathcal{B}} \to \mathbb{R}$ by $\mu(y) = \sum_a y_a$. Then,*

$$\min_{y \in (\mathcal{U}+u_0) \cap \mu^{-1}(t)} \|y\|_1 \leq 2 \min_{y \in \mathcal{U}+u_0} \|y\|_1 + |t|$$

Proof Define y^* by

$$y^* := \underset{y \in \mathcal{U}+u_0}{\operatorname{argmin}} \|y\|_1$$

Take any $\xi \in \mathcal{U} \cap \Delta\mathcal{B}$. Define $y^!$ by

$$y^! := y^* + (t - \mu(y^*))\,\xi$$

We have,

$$\begin{aligned}\mu\left(y^!\right) &= \mu\left(y^* + (t - \mu(y^*))\,\xi\right) \\ &= \mu(y^*) + (t - \mu(y^*))\,\mu(\xi) \\ &= \mu(y^*) + t - \mu(y^*) \\ &= t\end{aligned}$$

Hence, $y^! \in (\mathcal{U} + u_0) \cap \mu^{-1}(t)$. Moreover,

$$\begin{aligned}\left\|y^!\right\|_1 &= \|y^* + (t - \mu(y^*))\,\xi\|_1 \\ &\le \|y^*\|_1 + |t - \mu(y^*)| \cdot \|\xi\|_1 \\ &= \|y^*\|_1 + |t - \mu(y^*)| \\ &\le \|y^*\|_1 + |\mu(y^*)| + |t| \\ &\le 2\,\|y^*\|_1 + |t|\end{aligned}$$

■

We now demonstrate a bound on the norm on $\mathcal{W}$ in terms of the norms on the direct summands $\mathcal{W}_a$. It is achieved by repeatedly applying the previous lemma to construct the needed vector.

Lemma 45 *In the setting of Proposition 13, for any $x \in \mathcal{A}$ and $\theta \in \mathcal{H}$, $F_{x\theta}$ is onto. Moreover, for any $w \in \mathcal{W}$*

$$\|w\| \le 2 \sum_{\substack{i<n \\ a \in \mathcal{G}^i_-}} \|w_a\|$$

Here, the norms on the right hand side are induced by F_a and $\mathcal{H}_a$ in the usual way.

Proof In the following, we will use Lemma 33 implictly to justify the appearance of ℓ_1 norms. We use induction on n.

For $n = 1$, the sum on the right hand side has only one term which is clearly equal to $\|w\|$. Hence, the inequality holds.

Now, assume the claim for some $n \ge 1$ and let's show it for $n + 1$. Consider any $x \in \mathcal{A}$, $\theta \in \mathcal{H}$ and $w \in \mathcal{W}$. It is enough to show that there is some $y^* \in \mathbb{R}^{\mathcal{B}}$ (which may depend on x and θ) s.t.

$$F_{x\theta} y^* = w$$

$$\|y^*\|_1 \leq 2 \sum_{\substack{i<n \\ a \in \mathcal{G}^i_-}} \|w_a\|$$

Let $\mathcal{Z}^n$, $\mathcal{W}^n$, F^n and $\mathcal{H}^n$ be defined in the same way as $\mathcal{Z}$, $\mathcal{W}$, F and $\mathcal{H}$ respectively, except for n instead of $n+1$. We have

$$\mathcal{W} \cong \mathcal{W}^n \oplus \bigoplus_{a \in \mathcal{G}^n_-} \mathcal{W}_a$$

Denote by $w_{:n}$ the projection of w to $\mathcal{W}^n$. By definition of the norm on $\mathcal{W}^n$, we can choose some $\tilde{y} \in \mathbb{R}^{\mathcal{G}^n_-}$ s.t.

$$F^n_{x\theta}\tilde{y} = w_{:n}$$

$$\|\tilde{y}\|_1 \leq \|w_{:n}\|$$

For each $a \in \mathcal{G}^n_-$, apply Lemma 44 with

$$\mathcal{U} := \ker F_{ax\theta_a}$$

$$u_0 \in F^{-1}_{ax\theta_a}(w_a)$$

$$t := \tilde{y}_a$$

Notice that

$$\mathcal{U} + u_0 = F^{-1}_{ax\theta_a}(w_a)$$

This produces $u^a \in \mathbb{R}^{\mathcal{G}_n}$ s.t.

$$F_{ax\theta_a}u^a = w_a$$

$$\mu\left(u^a\right) = \tilde{y}_a$$

$$\begin{aligned}\|u^a\|_1 &\leq 2 \min_{y \in \mathbb{R}^{\mathcal{G}_n} : F_{ax\theta_a}y = w_a} \|y\|_1 + |\tilde{y}_a| \\ &\leq 2\|w_a\| + |\tilde{y}_a|\end{aligned}$$

Define $y^* \in \mathbb{R}^{\mathcal{B}}$ by $y^*_{ab} := u^a_b$ for any $a \in \mathcal{G}^n_-$ and $b \in \mathcal{G}_n$. Let's check that $F_{x\theta}y^* = w$.
For any $i < n$, define

$$\mathcal{G}^i_{(+)} := \prod_{i<j<n} \mathcal{G}_j$$

For any $i < n$ and $a \in \mathcal{G}_-^i$, we have

$$\begin{aligned}
(y^*)_b^a &= \sum_{c \in \mathcal{G}_+^i} y^*_{abc} \\
&= \sum_{d \in \mathcal{G}_{(+)}^i} \sum_{e \in \mathcal{G}_n} y^*_{abde} \\
&= \sum_{d \in \mathcal{G}_{(+)}^i} \sum_{e \in \mathcal{G}_n} u_e^{abd} \\
&= \sum_{d \in \mathcal{G}_{(+)}^i} \mu \left(u_e^{abd} \right) \\
&= \sum_{d \in \mathcal{G}_{(+)}^i} \tilde{y}_{abd} \\
&= \tilde{y}_b^a
\end{aligned}$$

We got $(y^*)^a = \tilde{y}^a$. It follows that,

$$\begin{aligned}
(F_{x\theta} y^*)_a &= F_{ax\theta_a} (y^*)^a \\
&= F_{ax\theta_a} \tilde{y}^a \\
&= (F_{x\theta}^n \tilde{y})_a \\
&= (w_{:n})_a \\
&= w_a
\end{aligned}$$

Now, consider any $a \in \mathcal{G}_-^n$. We have,

$$\begin{aligned}
(y^*)_b^a &= y^*_{ab} \\
&= u_b^a
\end{aligned}$$

Hence, $(y^*)^a = u^a$. It follows that,

$$\begin{aligned}
(F_{x\theta} y^*)_a &= F_{ax\theta_a} (y^*)^a \\
&= F_{ax\theta_a} u^a \\
&= w_a
\end{aligned}$$

It remains only to bound the norm of y^*. We have,

$$\begin{aligned}
\|y^*\|_1 &= \sum_{a\in\mathcal{B}} |y^*_a| \\
&= \sum_{b\in\mathcal{G}^b_-} \sum_{c\in\mathcal{G}_n} |y^*_{bc}| \\
&= \sum_{b\in\mathcal{G}^b_-} \sum_{c\in\mathcal{G}_n} \left|u^b_c\right| \\
&= \sum_{b\in\mathcal{G}^b_-} \left\|u^b\right\|_1 \\
&\leq \sum_{b\in\mathcal{G}^b_-} (2\,\|w_b\| + |\tilde{y}_b|) \\
&= 2\sum_{b\in\mathcal{G}^b_-} \|w_b\| + \sum_{b\in\mathcal{G}^b_-} |\tilde{y}_b| \\
&= 2\sum_{b\in\mathcal{G}^b_-} \|w_b\| + \|\tilde{y}\|_1 \\
&\leq 2\sum_{b\in\mathcal{G}^b_-} \|w_b\| + \|w_{:n}\|
\end{aligned}$$

Applying the induction hypothesis to the second term on the right hand side, we get

$$\begin{aligned}
\|y^*\|_1 &\leq 2\sum_{a\in\mathcal{G}^a_-} \|w_a\| + 2\sum_{\substack{i<n \\ a\in\mathcal{G}^i_-}} \|w_a\| \\
&= 2\sum_{\substack{i<n+1 \\ a\in\mathcal{G}^i_-}} \|w_a\|
\end{aligned}$$

■

Equipped with the previous lemma, the proof of Proposition 13 is straightforward.

Proof [Proof of Proposition 13]Once again, we will use Lemma 33 implictly. First, observe that for any $y \in \mathbb{R}^{\mathcal{B}}$ and $i < n$, we have

$$\begin{aligned}
\sum_{a\in\mathcal{G}^i_-} \|y^a\|_1 &= \sum_{a\in\mathcal{G}^i_-} \sum_{b\in\mathcal{G}_i} |y^a_b| \\
&= \sum_{a\in\mathcal{G}^i_-} \sum_{b\in\mathcal{G}_i} \left| \sum_{c\in\mathcal{G}^i_+} y_{abc} \right| \\
&\leq \sum_{a\in\mathcal{G}^i_-} \sum_{b\in\mathcal{G}_i} \sum_{c\in\mathcal{G}^i_+} |y_{abc}| \\
&= \|y\|_1
\end{aligned}$$

Now, let's analyze $R(\mathcal{H}, F)$:

$$\begin{aligned}
R\left(\mathcal{H}, F\right) &= \max_{\theta\in\mathcal{H}} \|\theta\| \\
&= \max_{\theta\in\mathcal{H}} \max_{x\in\mathcal{A}} \|F_{x\theta}\| \\
&= \max_{\substack{\theta\in\mathcal{H}\\ x\in\mathcal{A}}} \max_{y\in\mathbb{R}^{\mathcal{B}}:\|y\|=1} \|F_{x\theta} y\|
\end{aligned}$$

Applying Lemma 45 to the right hand side, we get

$$\begin{aligned}
R\left(\mathcal{H}, F\right) &\leq 2 \max_{\substack{\theta\in\mathcal{H}\\ x\in\mathcal{A}}} \max_{y\in\mathbb{R}^{\mathcal{B}}:\|y\|=1} \sum_{\substack{i<n\\ a\in\mathcal{G}^i_-}} \|(F_{x\theta} y)_a\| \\
&= 2 \max_{\substack{\theta\in\mathcal{H}\\ x\in\mathcal{A}}} \max_{y\in\mathbb{R}^{\mathcal{B}}:\|y\|=1} \sum_{\substack{i<n\\ a\in\mathcal{G}^i_-}} \|F_{ax\theta} y^a\| \\
&\leq 2 \max_{y\in\mathbb{R}^{\mathcal{B}}:\|y\|=1} \max_{\substack{\theta\in\mathcal{H}\\ x\in\mathcal{A}}} \sum_{\substack{i<n\\ a\in\mathcal{G}^i_-}} \|F_{ax\theta}\| \cdot \|y^a\|_1 \\
&\leq 2 \max_{y\in\mathbb{R}^{\mathcal{B}}:\|y\|=1} \sum_{\substack{i<n\\ a\in\mathcal{G}^i_-}} \max_{\substack{\theta\in\mathcal{H}\\ x\in\mathcal{A}}} \|F_{ax\theta}\| \cdot \|y^a\|_1 \\
&= 2 \max_{y\in\mathbb{R}^{\mathcal{B}}:\|y\|=1} \sum_{\substack{i<n\\ a\in\mathcal{G}^i_-}} R\left(\mathcal{H}_a, F_a\right) \|y^a\|_1 \\
&\leq 2 \left(\max_{\substack{i<n\\ a\in\mathcal{G}^i_-}} R\left(\mathcal{H}_a, F_a\right) \right) \max_{y\in\mathbb{R}^{\mathcal{B}}:\|y\|=1} \sum_{\substack{i<n\\ a\in\mathcal{G}^i_-}} \|y^a\|_1 \\
&\leq 2 \left(\max_{\substack{i<n\\ a\in\mathcal{G}^i_-}} R\left(\mathcal{H}_a, F_a\right) \right) \sum_{i<n} \max_{y\in\mathbb{R}^{\mathcal{B}}:\|y\|=1} \sum_{a\in\mathcal{G}^i_-} \|y^a\|_1 \\
&\leq 2 \left(\max_{\substack{i<n\\ a\in\mathcal{G}^i_-}} R\left(\mathcal{H}_a, F_a\right) \right) \sum_{i<n} \max_{y\in\mathbb{R}^{\mathcal{B}}:\|y\|=1} \|y\|_1 \\
&= 2 \left(\max_{\substack{i<n\\ a\in\mathcal{G}^i_-}} R\left(\mathcal{H}_a, F_a\right) \right) \sum_{i<n} 1 \\
&= 2n \max_{\substack{i<n\\ a\in\mathcal{G}^i_-}} R\left(\mathcal{H}_a, F_a\right)
\end{aligned}$$

■

References

T. Augustin, F.P.A. Coolen, G. de Cooman, and M.C.M. Troffaes. *Introduction to Imprecise Probabilities*. Wiley Series in Probability and Statistics. Wiley, 2014. ISBN 9781118763148. URL `https://books.google.co.il/books?id=sntcAwAAQBAJ`.

Gerald Beer. *Topologies on Closed and Closed Convex Sets*, volume 268 of *Mathematics and Its Applications*. Springer Netherlands, 1993. ISBN 9789401581493. doi: 10.1007/978-94-015-8149-3. URL `http://dx.doi.org/10.1007/978-94-015-8149-3`.

Zixiang Chen, Dongruo Zhou, and Quanquan Gu. Almost optimal algorithms for two-player zero-sum linear mixture markov games. In Sanjoy Dasgupta and Nika Haghtalab, editors, *International Conference on Algorithmic Learning Theory, 29-1 April 2022, Paris, France*, volume 167 of *Proceedings of Machine Learning Research*, pages 227–261. PMLR, 2022. URL `https://proceedings.mlr.press/v167/chen22d.html`.

Varsha Dani, Thomas P. Hayes, and Sham M. Kakade. Stochastic linear optimization under bandit feedback. In Rocco A. Servedio and Tong Zhang, editors, *21st Annual Conference on Learning Theory - COLT 2008, Helsinki, Finland, July 9-12, 2008*, pages 355–366. Omnipress, 2008. URL `http://colt2008.cs.helsinki.fi/papers/80-Dani.pdf`.

Uriel Feige. Introduction to semirandom models. In Tim Roughgarden, editor, *Beyond the Worst-Case Analysis of Algorithms*, pages 189–211. Cambridge University Press, 2020. doi: 10.1017/9781108637435.013. URL `https://doi.org/10.1017/9781108637435.013`.

A. A. Giannopoulos. A note on the banach-mazur distance to the cube. In J. Lindenstrauss and V. Milman, editors, *Geometric Aspects of Functional Analysis*, pages 67–73, Basel, 1995. Birkhäuser Basel. ISBN 978-3-0348-9090-8.

G. Giorgi and T.H. Kjeldsen. *Traces and Emergence of Nonlinear Programming*. SpringerLink : Bücher. Springer Basel, 2013. ISBN 9783034804394. URL `https://books.google.co.il/books?id=q4PBBAAAQBAJ`.

Achim Klenke. *Probability Theory: A Comprehensive Course*. Springer International Publishing, 2020. ISBN 9783030564025. doi: 10.1007/978-3-030-56402-5. URL `http://dx.doi.org/10.1007/978-3-030-56402-5`.

Vanessa Kosoy. Imprecise multi-armed bandits. Master's thesis, Hebrew University of Jerusalem, 2024. URL `https://arxiv.org/abs/2405.05673`.

T. Lattimore and C. Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2020. ISBN 9781108486828. URL `https://books.google.co.il/books?id=bbjpDwAAQBAJ`.

Ankur Moitra, William Perry, and Alexander S. Wein. How robust are reconstruction thresholds for community detection? In Daniel Wichs and Yishay Mansour, editors, *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 828–841. ACM, 2016. doi: 10.1145/2897518.2897573. URL `https://doi.org/10.1145/2897518.2897573`.

Brendan O'Donoghue, Tor Lattimore, and Ian Osband. Matrix games with bandit feedback. In Cassio P. de Campos, Marloes H. Maathuis, and Erik Quaeghebeur, editors, *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence, UAI 2021, Virtual Event, 27-30 July 2021*, volume 161 of *Proceedings of Machine Learning Research*, pages 279–289. AUAI Press, 2021. URL https://proceedings.mlr.press/v161/o-donoghue21a.html.

Daniel Russo and Benjamin Van Roy. Eluder dimension and the sample complexity of optimistic exploration. In C.J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc., 2013. URL https://proceedings.neurips.cc/paper/2013/file/41bfd20a38bb1b0bec75acf0845530a7-Paper.pdf.

Jacob Steinhardt. Does robustness imply tractability? A lower bound for planted clique in the semi-random model. *Electron. Colloquium Comput. Complex.*, TR17-069, 2017. URL https://eccc.weizmann.ac.il/report/2017/069.

Yi Tian, Yuanhao Wang, Tiancheng Yu, and Suvrit Sra. Online learning in unknown markov games. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 10279–10288. PMLR, 18–24 Jul 2021. URL https://proceedings.mlr.press/v139/tian21b.html.