Pure Exploration with Infinite Answers

Riccardo Poiani
Bocconi University

Martino Bernasconi
Bocconi University

Andrea Celli
Bocconi University

Abstract

We study pure exploration problems where the set of correct answers is possibly infinite, *e.g.*, the regression of any continuous function of the means of the bandit. We derive an instance-dependent lower bound for these problems. By analyzing it, we discuss why existing methods (*i.e.*, Sticky Track-and-Stop) for finite answer problems fail at being asymptotically optimal in this more general setting. Finally, we present a framework, Sticky-Sequence Track-and-Stop, which generalizes both Track-and-Stop and Sticky Track-and-Stop, and that enjoys asymptotic optimality. Due to its generality, our analysis also highlights special cases where existing methods enjoy optimality.

1 INTRODUCTION

In *pure exploration* problems, an agent sequentially interacts with a set of $K \in \mathbb{N}$ probability distributions denoted by $\boldsymbol{\nu} = (\nu_k)_{k\in[K]} \in \mathcal{Q}$ modeling the outcome of K different experiments, where $\mathcal{Q}$ is an arbitrary set of problems. The main goal of the agent is answering a given question about these distributions as efficiently as possible, *i.e.*, using the least possible amount of samples. Let $\mathcal{X}$ be an answer space for the question at hand; then, for each possible $\boldsymbol{\nu} \in \mathcal{Q}$, a set-valued function (a.k.a. *correspondence* [Aubin, 1999]) maps each possible instance $\boldsymbol{\nu}$ to a set of correct answers $\mathcal{X}^\star(\boldsymbol{\nu}) \subseteq \mathcal{X}$. The agent is then given a maximum risk parameter $\delta \in (0,1)$ and has to return a correct answer $x \in \mathcal{X}^\star(\boldsymbol{\nu})$ with probability at least $1-\delta$, while minimizing the number of interactions.

This framework models a broad range of settings, with the most extensively studied being the *Best-Arm Identification* (BAI) problem [Even-Dar et al., 2002]. Here, the answer space is $\mathcal{X} = [K]$ and the unique correct answer is described by the single-valued correspondence $\mathcal{X}^\star(\boldsymbol{\nu}) = \operatorname{argmax}_{k\in[K]} \mu_k$, where $\boldsymbol{\mu} = (\mu_k)_{k\in[K]}$ denotes the means of the distributions in $\boldsymbol{\nu}$.[1] In the seminal work by Garivier and Kaufmann [2016], the authors derived an information-theoretic lower bound showing that, in the *unstructured* bandit setting, any algorithm requires at least a certain number of samples in order to identify the best arm with high probability. Furthermore, the authors proposed the Track-and-Stop (TaS) algorithm, which achieves optimal sample complexity rates in the high confidence regime of $\delta \to 0$. The key idea behind TaS is to exploit *oracle weights*, which are probability distributions over arms that represents the optimal sampling strategy for an algorithm with full knowledge of the instance $\boldsymbol{\nu}$. TaS mimics this oracle algorithm by tracking the oracle weights of an empirical estimate of the instance $\boldsymbol{\mu}$. Subsequent work has extended these asymptotic optimality results to BAI problems with additional structure on the instance means $\boldsymbol{\mu}$ [Moulos, 2019, Degenne et al., 2020a, Kocák and Garivier, 2020, Jourdan et al., 2021, Russac et al., 2021, Vannella et al., 2023, Carlsson et al., 2024, Poiani et al., 2024, 2025, Russo et al., 2025, Tuynman and Degenne, 2025, Lazzaro and Pike-Burke, 2025]. More generally, several works have shown how to leverage these techniques to build asymptotically optimal algorithms for arbitrarily structured problems where there is a *single* correct answer [Degenne et al., 2019, Ménard, 2019, Juneja and Krishnasamy, 2019, Wang et al., 2021, Kaufmann and Koolen, 2021]. For instance, asymptotically optimal results are available for the broad class of *sequential partition identification problems*, where each instance $\boldsymbol{\nu}$ belongs to an element of a partition of the set of instances $\mathcal{Q}$, and the goal lies in identifying the index of the partition that $\boldsymbol{\nu}$ belongs to. A key aspect of most of these studies is that they rely on the *well-behaved* nature of oracle weights as functions of the instance.

These properties no longer hold in the broader setting of problems with *multiple correct answers*, where the

Proceedings of the 29th International Conference on Artificial Intelligence and Statistics (AISTATS) 2026, Tangier, Morocco. PMLR: Volume 300. Copyright 2026 by the author(s).

[1] For any $n \in \mathbb{N}$ we denote by $[n]$ the set $\{1, \dots, n\}$.

correspondence $\mathcal{X}^\star(\boldsymbol{\mu})$ denoting the set of correct answers is no longer single-valued. Asymptotic optimality for such problems has been studied by Degenne and Koolen [2019] in the case in which the set of possible answers $\mathcal{X}$ is finite.[2] Specifically, the statistical lower bound is now expressed as a minimum over all the multiple correct answers of the lower bound for single-answer problems. Let $\mathcal{X}_F(\boldsymbol{\nu}) \subseteq \mathcal{X}^\star(\boldsymbol{\nu})$ denote the subset of correct answers that attain this minimum. Intuitively, when there are multiple correct answers, some answers might be statistically easier to identify than others. $\mathcal{X}_F(\boldsymbol{\nu})$ is precisely the set of the "easiest" correct answers. However, the presence of multiple answers in $\mathcal{X}_F(\boldsymbol{\nu})$ introduces topological challenges that hinder the direct application of the Track-and-Stop algorithm. Degenne and Koolen [2019] solve this issue by introducing the Sticky-Track-and-Stop (Stiky-TaS) algorithm, which first identifies a statistically convenient correct answer, *i.e.*, one that belongs to $\mathcal{X}_F(\boldsymbol{\nu})$, and then sticks to it by tracking its corresponding oracle weights (which, for a fixed answer, exhibit the "nice" properties required to prove asymptotic optimality). Crucially, both the way Sticky-TaS selects an answer and its ability to stick to it heavily depend on the fact that $|\mathcal{X}|$ is finite.

In this work, we drop the assumption that $|\mathcal{X}|$ is finite and study the more general setting in which the answer set $\mathcal{X}$ may be infinite. Our analysis adopts an asymptotically optimal perspective, focusing on the challenges that arise when the set of correct answers $\mathcal{X}^\star(\boldsymbol{\mu})$ is infinite. This model captures fundamental applications that are currently underexplored and not yet fully understood in the bandit literature, such as the problem of regressing a continuous function of the bandit means. To give a concrete application example, consider a pricing problem where each arm corresponds to posting a certain price, and the observed reward is a Bernoulli variable indicating whether the item was purchased. A company may wish to estimate, up to accuracy $\epsilon > 0$, the revenue associated with the optimal price, or the maximum revenue gap that exists between the optimal and the worst price. In these cases, both the answer space and the space of correct answers are infinite sets.

1.1 Contributions

We focus on what we call *regular* pure exploration problems. Intuitively, these are problems in which alternative models for an answer $x \in \mathcal{X}$ are "stable" for nearby answers x' (*i.e.*, belonging to some neighborhood of x). This definition (see Section 3.1 for the precise statement) imposes only minimal and natural requirements. Indeed, it encompasses all cases where the correspondence $\boldsymbol{\mu} \rightrightarrows \mathcal{X}^\star(\boldsymbol{\mu})$ is continuous (Theorem 1). This continuity property is satisfied in fundamental problems such as the problem of regressing an arbitrary continuous function of $\boldsymbol{\mu}$.

After introducing regular pure exploration problems, we present an asymptotic lower bound on the number of samples that are required to identify a correct answer in $\mathcal{X}^\star(\boldsymbol{\mu})$ (Theorem 2). Then, in Section 4, we analyze the properties of this lower bound. In particular, we study continuity properties of the oracle weights, the lower bound itself, and the mapping $\mathcal{X}_F(\boldsymbol{\mu})$, and we analyze their algorithmic implications for infinite-answer problems. We argue that the presence of infinite answers makes it impossible to select and track the empirical oracle weights for a *single* correct answer in $\mathcal{X}_F(\boldsymbol{\nu})$. This undermines the core argument behind asymptotic optimality of Sticky-TaS.

We address this challenge in Section 5. In particular, we show that it is not necessary to select and stick to a single correct answer. Instead, it suffices to track a sequence of empirical oracle weights associated with a sequence of answers that converges to some (potentially unknown a priori) correct answer in $\mathcal{X}_F(\boldsymbol{\nu})$. Building on this, we introduce a general framework (Sticky-Sequence Track-and-Stop) which, when equipped with a method for selecting a converging sequence of answers, achieves asymptotic optimality guarantees (Theorem 3).

The main challenge here is constructing a converging sequence of answers by only exploiting a sequence of sets that converges to the unknown set $\mathcal{X}_F(\boldsymbol{\nu})$. To present challenges and connections with existing algorithms (*i.e.*, TaS and Sticky-TaS), we show how this can be achieved according to different topological properties of $\mathcal{X}$ and $\mathcal{X}_F(\boldsymbol{\nu})$, which we group in four main scenarios. **(i)** If $\mathcal{X}_F(\boldsymbol{\nu})$ is single-valued for all problems in $\mathcal{Q}$, then both TaS and Sticky-TaS already implement a converging sequence of answers, thereby achieving optimality. **(ii)** When $\mathcal{X} \subset \mathbb{R}$, this property is lost by TaS, but the total order on the reals ensures that the sequence of answers selected by Sticky-TaS converges to some $x \in \mathcal{X}_F(\boldsymbol{\mu})$. **(iii)** If $|\mathcal{X}_F(\boldsymbol{\nu})|$ is finite but, for example, $\mathcal{X} \subset \mathbb{R}^2$, then neither TaS nor Sticky-TaS guarantees the convergence property. However, this can be ensured by a simple rule that selects the next answer as the closest to the previous one within a suitable confidence region. **(iv)** In the general case where the only information available is that $\mathcal{X} \subset \mathbb{R}^d$, we propose an algorithm that progressively discretizes the answer space while guiding the selection of answers according to the history of the previously selected ones.

[2] This formulation also models problems like ϵ-best arm identification, further explored in Kocák and Garivier [2021], Jourdan and Degenne [2022], Jourdan et al. [2023].

Additional Related Works In pure exploration literature, to the best of our knowledge, there exists two works that deals with infinite answer problems *i.e.*, Deep et al. [2024], Osogami et al. [2025]. In particular, Deep et al. [2024] focus on unstructured $K = 2$ bandit problems with the goal of estimating an interval of width at most ϵ that contains $\mu_1 - \mu_2$ with high probability. Osogami et al. [2025], instead, study the problem of estimating the value of the optimal arm, up to an accuracy ϵ, in unstructured bandit settings. Our work is more general in the sense that we deal with correct answer correspondences that comprehend the ones considered in Deep et al. [2024], Osogami et al. [2025]. Furthermore, structured bandits are also allowed within our framework.

2 PRELIMINARIES

Mathematical Background We denote by Δ_n the n-dimensional simplex. Given $\mathcal{X} \subseteq \mathbb{R}^d$ and $x \in \mathcal{X}$, we denote by $\mathcal{B}_\rho(x) = \{x' \in \mathcal{X} : \|x - x'\| \leq \rho\}$ the ball of radius ρ around x and for a set $A \subset \mathcal{X}$ we denote by $\mathcal{B}_\rho(A)$ the union of $\mathcal{B}_\rho(x)$ over all $x \in A$. Given a set $\mathcal{X}$, we denote by $\mathrm{cl}(\mathcal{X})$ its closure. Now, let $\mathcal{X} \subseteq \mathbb{R}^{n_x}$ and $\mathcal{Y} \subseteq \mathbb{R}^{n_y}$, we denote by $C : \mathcal{X} \rightrightarrows \mathcal{Y}$ a set-valued function (i.e., *correspondence*) that maps each element $x \in \mathcal{X}$ to a (non-empty) subset $C(x) \subset \mathcal{Y}$ [Aubin, 1999]. The correspondence C is *upper hemicontinuous* if, for all $x \in \mathcal{X}$, and for every open set $\mathcal{V} \subset \mathcal{Y}$ such that $C(x) \subset \mathcal{V}$, there exists a neighbourhood $\mathcal{U}$ of x such that $C(x')$ is a subset of $\mathcal{V}$ for all $x' \in \mathcal{U}$. Furthermore, C is *lower hemicontinuous* if, for all $x \in \mathcal{X}$, and for every open set $\mathcal{V} \subset \mathcal{Y}$ such that $C(x) \cap \mathcal{V} \neq \emptyset$, there exists a neighbourhood $\mathcal{U}$ of x such that $C(x') \cap \mathcal{V} \neq \emptyset$ for all $x' \in \mathcal{U}$. Finally, a correspondence C is *continuous* if it is upper and lower hemicontinuous.

Learning Model The learner has $K \in \mathbb{N}$ possible choices, each associated with a probability distribution ν_k over $\mathbb{R}$. We denote by $\boldsymbol{\nu} = (\nu_k)_{k\in[K]}$ the vector of distributions which we refer to as the bandit model. Let $\boldsymbol{\mu} = (\mu_k)_{k\in[K]}$ be the vector of means of the distributions ν_k, where each $\mu_k = \mathbb{E}_{R\sim\nu_k}[R]$ is the expected value under distribution ν_k. In this work, we consider distributions that belong to a canonical exponential family [Cappé et al., 2013].[3] Conveniently, since these distributions are fully characterized by their means, we may, with a slight abuse of notation, refer to $\boldsymbol{\mu}$ as the bandit model. We denote by Θ the interval defining the possible means for any arm μ. We make the standard assumption that the exponential family is regular and bounded, meaning that Θ is a closed interval strictly contained in an open one (see, *e.g.*, Degenne et al. [2019], Poiani et al. [2025]). Furthermore, to represent (possible) additional structure on the bandit model, we assume knowledge of a set $\mathcal{M} \subseteq \Theta^K$ defining the set of admissible bandits models that the learner could face, *i.e.*, $\boldsymbol{\mu} \in \mathcal{M}$. We consider a possibly infinite answer space $\mathcal{X} \subseteq \mathbb{R}^d$. For each bandit model $\boldsymbol{\mu} \in \Theta^K$, the set of correct answers for $\boldsymbol{\mu}$ is represented by a correspondence $\mathcal{X}^\star : \Theta^K \rightrightarrows \mathcal{X}$. The learner interacts repeatedly with the bandit model. During each round $t \in \mathbb{N}$, it selects an action $A_t \in [K]$ and observes an outcome $R_t \sim \nu_{A_t}$. Let $\mathcal{F}_t = \sigma((A_s, R_s)_{s=1}^t)$ be the σ-field generated by the observations up to time t. A learning algorithm takes in input a risk parameter $\delta \in (0, 1)$ and is composed of (i) a $\mathcal{F}_{t-1}$-measurable sampling rule that selects the next action $A_t \in [K]$, (ii) a stopping rule τ_δ, which is a stopping time with respect to $(\mathcal{F}_t)_{t\in\mathbb{N}}$, and (iii) a $\mathcal{F}_{\tau_\delta}$-measurable decision rule that selects a final decision $\hat{x}_{\tau_\delta} \in \mathcal{X}$. We say that an algorithm is *δ-correct* if $\mathbb{P}_{\boldsymbol{\mu}}(\hat{x}_{\tau_\delta} \notin \mathcal{X}^\star(\boldsymbol{\mu})) \leq \delta$ for all $\boldsymbol{\mu} \in \mathcal{M}$. In words, δ-correct algorithms provide, for each $\boldsymbol{\mu} \in \mathcal{M}$, an answer $\hat{x}_{\tau_\delta} \in \mathcal{X}^\star(\boldsymbol{\mu})$ among the correct ones with probability at least $1 - \delta$. Among the class of δ-correct algorithms, we look for those minimizing the expected stopping time, that is, $\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta] = \sum_{k\in[K]} \mathbb{E}_{\boldsymbol{\mu}}[N_k(\tau_\delta)]$, where $N_k(t)$ denotes the (random) number of samples collected for action $k \in [K]$ up to time t. In the following, we will use $\boldsymbol{N}(t)$ to denote the vector $(N_1(t), \dots, N_k(t))$ and $\hat{\boldsymbol{\mu}}(t)$ to denote the empirical estimate of $\boldsymbol{\mu}$ at time t, *i.e.*, $\hat{\mu}_k(t) = N_k(t)^{-1} \sum_{s=1}^t R_s \mathbf{1}\{A_t = k\}$.

Alternative Models For each $x \in \mathcal{X}$, the set of *alternative models* $\neg x$ is defined as $\neg x = \{\boldsymbol{\lambda} \in \mathcal{M} : x \notin \mathcal{X}^\star(\boldsymbol{\lambda})\}$ (see, *e.g.*, Degenne et al. [2019], Degenne and Koolen [2019]). In words, $\neg x$ is the set that contains all the bandit models $\boldsymbol{\lambda}$ for which x is *not* a correct answer for $\boldsymbol{\lambda}$.[4] We generalize this concept to any subset of answers $\widetilde{\mathcal{X}} \subseteq \mathcal{X}$. Specifically, $\neg\widetilde{\mathcal{X}} = \{\boldsymbol{\lambda} \in \mathcal{M} : \forall x \in \widetilde{\mathcal{X}} : x \notin \mathcal{X}^\star(\boldsymbol{\lambda})\}$. The set $\neg\widetilde{\mathcal{X}}$ extends the notion of alternative models to any arbitrary collection of answers, as it requires that *each* answer $x \in \widetilde{\mathcal{X}}$ is not correct for $\boldsymbol{\lambda}$. It directly follows that $\neg\widetilde{\mathcal{X}} \subseteq \neg x$ for any $\widetilde{\mathcal{X}}$ such that $x \in \widetilde{\mathcal{X}}$. This generalization plays a crucial role in defining *regular pure exploration problems*, which we introduce in Section 3.1.

[3] These distributions include, *e.g.*, Gaussian distributions with known variance and Bernoulli distributions. We refer the interested reader to Appendix H.2 for additional details on canonical exponential families. For concentration purposes, we also require that these distributions are sub-Gaussian.

[4] W.l.o.g., we assume that $\neg x \neq \emptyset$ for all $x \in \mathcal{X}$. Indeed, if there exists $\bar{x} \in \mathcal{X}$ such that $\neg\bar{x} = \emptyset$, then a δ-correct algorithm can trivially return $\bar{x}$ for any $\boldsymbol{\mu} \in \mathcal{M}$ without even interacting with the environment.

Divergences Finally, we introduce some divergences that are commonly used in pure exploration problems (see *e.g.*, Degenne et al. [2019], Degenne and Koolen [2019]). Intuitively, as we will see in the next section, they are helpful in statistically identifying the correct answers. Let $d(p, q)$ be the KL divergence between distributions with means p and q, respectively. Then, for $\boldsymbol{\mu} \in \Theta^K, \Lambda \subseteq \mathcal{M}$ and $\boldsymbol{\omega} \in \mathbb{R}^K$, we define the following:

$$D(\boldsymbol{\mu}, \boldsymbol{\omega}, \Lambda) = \inf_{\boldsymbol{\lambda} \in \Lambda} \sum_{k \in [K]} \omega_k d(\mu_k, \lambda_k)$$
$$D(\boldsymbol{\mu}, \Lambda) = \sup_{\boldsymbol{\omega} \in \Delta_K} D(\boldsymbol{\mu}, \boldsymbol{\omega}, \Lambda)$$
$$D(\boldsymbol{\mu}) = \sup_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \neg x).$$

Let us introduce $\mathcal{X}_F(\boldsymbol{\mu}) = \operatorname{argmax}_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \neg x)$. We will later see that $\mathcal{X}_F(\boldsymbol{\mu})$ represents the subset of correct answers for $\boldsymbol{\mu}$ that are the "easiest" to identify. Moreover, for any $\boldsymbol{\mu} \in \Theta^K$ and any set $\Lambda \subseteq \mathcal{M}$, we denote by $\boldsymbol{\omega}^*(\boldsymbol{\mu}, \Lambda)$ the argmax over $\boldsymbol{\omega}$ of $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \Lambda)$. Finally, $\boldsymbol{\omega}^*(\boldsymbol{\mu}) = \bigcup_{x \in \mathcal{X}_F(\boldsymbol{\mu})} \boldsymbol{\omega}^*(\boldsymbol{\mu}, \neg x)$ denotes the oracle weights for $\boldsymbol{\mu}$.

3 LOWER BOUND

3.1 Regular Pure Exploration Problems

We now define the class of *regular pure exploration problems*, which are characterized by a set of regularity assumptions detailed below. All the problems considered in this paper belong to this class.

Assumption 1 (Compactness)**.** *$\mathcal{X}$ is compact and $\boldsymbol{\mu} \rightrightarrows \mathcal{X}^\star(\boldsymbol{\mu})$ is compact-valued.*

Assumption 2 (Identifiability)**.** *For all $\boldsymbol{\mu} \in \mathcal{M}$, there exists $\bar{x} \in \mathcal{X}^\star(\boldsymbol{\mu})$ such that $\boldsymbol{\mu} \notin \operatorname{cl}(\neg \bar{x})$.*

Assumption 3 (Continuity of $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg \mathcal{B}_\rho(x))$ to $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$)**.** *For all sufficiently small $\epsilon > 0$, there exists $\rho > 0$ such that, for all $\boldsymbol{\mu} \in \Theta^K$, $\boldsymbol{\omega} \in \Delta_K$, $x \in \mathcal{X}$, it holds that $\neg \mathcal{B}_\rho(x) \neq \emptyset$ and $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg \mathcal{B}_\rho(x)) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x) \leq \epsilon$.*

Assumption 1 imposes mild regularity conditions, namely compactness, on both the answer space and the correct answer correspondence $\mathcal{X}^\star(\boldsymbol{\mu})$. Assumption 2, instead, is necessary for learnability. When this assumption does not hold, the sample complexity is infinite, even in settings with a finite number of possible answers. This follows directly from the lower bound by Degenne and Koolen [2019] (see Appendix B.1 for further discussion). Finally, at first glance, Assumption 3 may appear to be a purely technical condition. However, as our analysis reveals, for a subset $\widetilde{\mathcal{X}} \subseteq \mathcal{X}^\star(\boldsymbol{\mu})$, the quantity $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg \widetilde{\mathcal{X}})$ can be related to the complexity of distinguishing $\boldsymbol{\mu}$ from all the models in $\mathcal{M}$ for which *none* of the answers in $\widetilde{\mathcal{X}}$ are correct. Intuitively, Assumption 3 implies that distinguishing $\boldsymbol{\mu}$ from $\neg x$ becomes arbitrarily similar to distinguishing $\boldsymbol{\mu}$ from $\neg \mathcal{B}_\rho(x)$ whenever ρ is small. In this sense, to demonstrate that Assumption 3 holds, it suffices to prove a form of "smoothness" in the alternative models when switching from $\neg x$ to $\neg \mathcal{B}_\rho(x)$, *i.e.*, for all $\boldsymbol{\lambda} \in \neg x$, there exists $\tilde{\boldsymbol{\lambda}} \in \neg \mathcal{B}_\rho(x)$ such that $\boldsymbol{\lambda} \approx \tilde{\boldsymbol{\lambda}}$ for $\rho \to 0$ (see Lemma 23). As one might expect, we can show that Assumption 3 holds whenever $\boldsymbol{\mu} \rightrightarrows \mathcal{X}^\star(\boldsymbol{\mu})$ is continuous (see Appendix B.2 for the proof).

Theorem 1 (Continuous Correspondence Implies Assumption 3)**.** *If $\boldsymbol{\mu} \rightrightarrows \mathcal{X}^\star(\boldsymbol{\mu})$ is continuous, and $\mathcal{M}$ and $\mathcal{X}$ are compact sets, then, Assumption 3 holds.*

Uniform vs Local Continuity It is important to observe that we require Assumption 3 to hold uniformly over $\Theta \times \Delta_K \times \mathcal{X}$. If Assumption 3 only required local continuity, then Theorem 1 would follow almost directly from the reasoning presented earlier. Showing uniform continuity is more challenging, as we need to show the existence of a ρ for which the conditions hold uniformly across all choices of $\boldsymbol{\lambda}$. At this point, one might wonder why we relied on Assumption 3 rather than directly assuming that $\mathcal{X}^\star(\boldsymbol{\mu})$ is continuous. The key point is that Assumption 3 allows us to fully generalize prior results for the finite-answers setting. Indeed, Degenne and Koolen [2019] allow $\mathcal{X}^\star(\boldsymbol{\mu})$ to be discontinuous in $\boldsymbol{\mu}$. Nevertheless, Assumption 3 always holds for problems with finite possible answers (see Appendix A.2), showing that we can properly generalize Degenne and Koolen [2019].

Examples We now give some examples of regular pure exploration problems. First, as anticipated above, we can deal with arbitrary finite answer problems (Appendix A.2). Secondly, given $\epsilon > 0$ and any continuous function $f : \Theta \to \mathcal{X}$ (*e.g.*, the maximum), we can consider the problem of estimating $f(\boldsymbol{\mu})$ up to an accuracy level ϵ.[5] In this case we have $\mathcal{X}^\star(\boldsymbol{\mu}) = \{x \in \mathcal{X} : \|f(\boldsymbol{\mu}) - x\|_\infty \leq \epsilon\}$. In Appendix A.1 we prove that, Assumptions 1 to 3 holds in these problems.[6] Furthermore, given two distinct learning problems defined by the correspondences $\mathcal{X}_1^\star$ and $\mathcal{X}_2^\star$ for which Assumptions 1 to 3 hold, one can prove that Assumptions 1 to 3 hold for the learning problem defined by the product correspondence $\mathcal{X}^\star(\boldsymbol{\mu}) = \{(x_1, x_2) : x_1 \in \mathcal{X}_1^\star(\boldsymbol{\mu}), x_2 \in \mathcal{X}_2^\star(\boldsymbol{\mu})\}$ (see Appendix A.3). Thus, we can combine arbitrary finite

[5] Note that, thanks to this result, we are able to show that regular exploration problems, under the assumption of a compact domain, encompasses the problem studied in Deep et al. [2024], Osogami et al. [2025].

[6] We introduced the problem using the ℓ_∞-norm, but other norms could also be considered.

answer problems with regression problems. Finally, we show in Appendix A.4 that the problem of learning an ϵ-Nash equilibrium in two player stochastic zero-sum games [*e.g.*, Zhou et al., 2017, Maiti et al., 2023] can be formulated as a regular pure exploration problem.

3.2 Sample Complexity Lower Bound

We now present the lower bound for infinite-answer problems, whose proof is deferred to Appendix C.

Theorem 2 (Lower Bound)**.** *For any $\boldsymbol{\mu} \in \mathcal{M}$, and any δ-correct algorithm it holds that:*

$$\liminf_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} \geq T^*(\boldsymbol{\mu}) = \frac{1}{D(\boldsymbol{\mu})}. \tag{1}$$

Theorem 2 provides an asymptotic lower bound on $\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]$ that holds for any δ-correct algorithm. By explicitly writing $D(\boldsymbol{\mu}) = \sup_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} \sup_{\boldsymbol{\omega} \in \Delta_k} \inf_{\boldsymbol{\lambda} \in \neg x} \sum_{k \in [K]} \omega_k d(\mu_k, \lambda_k)$ we see that the lower bound of Theorem 2 is expressed as a max-min game, where the max player chooses both a correct answer x within $\mathcal{X}^\star(\boldsymbol{\mu})$ and a strategy $\boldsymbol{\omega}$ over the arm space, and the min player chooses an alternative model $\boldsymbol{\lambda}$ for which x is not a correct answer. For this reason, answers in $\mathcal{X}_F(\boldsymbol{\mu})$ can be regarded as the statistically easiest correct answers to be identified (formally, we will prove that the sup over $\mathcal{X}^\star(\boldsymbol{\mu})$ is actually attained and therefore $\mathcal{X}_F(\boldsymbol{\mu})$ is well defined). Finally, we mention that Theorem 2 nicely generalizes the lower bound for multiple (but finite) correct answers of Degenne and Koolen [2019]. We note that the proof of the lower bound of Degenne and Koolen [2019] was explicitly using the fact that $\mathcal{X}^\star(\boldsymbol{\mu})$ is finite. Proving Theorem 2 thus required ad-hoc arguments to extend the result to our setting. For space constraints, we provide further details on this point in Appendix C.

4 ON REGULAR PURE EXPLORATION PROBLEMS

We now present properties of the divergences and discuss their implication for existing algorithms, *i.e.*, Track-and-Stop [Garivier and Kaufmann, 2016] and Sticky Track-and-Stop [Degenne and Koolen, 2019].

Lemma 1 (Continuity)**.** *The following holds:*

(i) The function $(\boldsymbol{\mu}, \boldsymbol{\omega}, x) \to D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is continuous over $\Theta^K \times \Delta_K \times \mathcal{X}$.

(ii) The function $(\boldsymbol{\mu}, x) \to D(\boldsymbol{\mu}, \neg x)$ is continuous over $\Theta^K \times \mathcal{X}$ and $(\boldsymbol{\mu}, x) \rightrightarrows \boldsymbol{\omega}^*(\boldsymbol{\mu}, \neg x)$ is upper hemicontinuous and compact-valued.

(iii) The function $(\boldsymbol{\mu}, \boldsymbol{\omega}) \to \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is continuous over $\Theta^K \times \Delta_K$ and $(\boldsymbol{\mu}, \boldsymbol{\omega}) \rightrightarrows \operatorname{argmax}_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is upper hemicontinuous and compact-valued.

(iv) The function $\boldsymbol{\mu} \to D(\boldsymbol{\mu})$ is continuous over Θ^K. Moreover, $\boldsymbol{\mu} \rightrightarrows \boldsymbol{\omega}^\star(\boldsymbol{\mu})$ and $\boldsymbol{\mu} \rightrightarrows \mathcal{X}_F(\boldsymbol{\mu})$ are upper hemicontinuous and compact-valued over $\mathcal{S}$.

It is well-known that results analogous to Lemma 1 play a crucial role in the design of optimal algorithms. For instance, Degenne and Koolen [2019] exploited similar results for the case of problems with finite sets of answers.

Although properties *(ii)* and *(iii)* play a crucial role in designing optimal algorithms, their derivation is relatively standard. In contrast, it is interesting to consider properties *(i)* and *(iv)*, for which we need to make at least three important considerations. First, proving point *(i)* requires novel arguments as, contrary to Degenne and Koolen [2019], we have to guarantee that $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is jointly continuous on the product space $\Theta^K \times \Delta_K \times \mathcal{X}$. The joint continuity is crucial here, since it is then used to prove all the other claims within Lemma 1. On a technical level, the main idea to prove *(i)* is combining the joint continuity of $(\boldsymbol{\mu}, \boldsymbol{\omega}) \to D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ for all $x \in \mathcal{X}$ (see [Degenne and Koolen, 2019, Theorem 4]) together with Assumption 3. To this end, it is important to highlight again that Assumption 3 holds uniformly across the domain. This allows us to prove joint continuity by using that $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is separately continuous in $\boldsymbol{\mu}$ and $\boldsymbol{\omega}$ for a fixed x. Second, it is interesting to highlight that, in point *(iv)*, the continuity of $\boldsymbol{\mu} \to D(\boldsymbol{\mu})$ and, more surprisingly, the upper hemicontinuity of $\mathcal{X}_F(\boldsymbol{\mu})$ do not require the continuity of $\boldsymbol{\mu} \rightrightarrows \mathcal{X}^\star(\boldsymbol{\mu})$ but only the joint continuity of $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$. Indeed, when studying $\max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \neg x)$, if $\mathcal{X}^\star(\boldsymbol{\mu})$ is not continuous, we cannot directly apply Berge's maximum theorem to prove the continuity of $D(\boldsymbol{\mu})$. However, we can observe that $\max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \neg x) = \max_{x \in \mathcal{X}} D(\boldsymbol{\mu}, \neg x)$ (Lemma 27) and then use Berge's theorem ($\mathcal{X}$ is a constant correspondence and thus trivially continuous).

4.1 Failure of Sticky Track-and-Stop with infinite answers

We now discuss why the Sticky-TaS algorithm [Degenne and Koolen, 2019] is not optimal in the infinite-answer setting. In particular, while conditions similar to those in Lemma 1 were sufficient to establish its optimality in the finite-answer case, they are no longer sufficient when the answer space is infinite. This discussion will underscore the fundamental differences between the finite and infinite-answer settings.

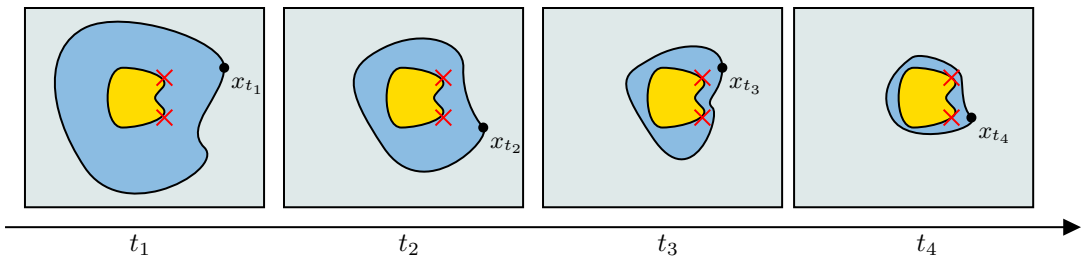

Figure 1: Even though the sets $\mathcal{X}_t$ (in **blue**) are progressively shrinking toward $\mathcal{X}_F(\boldsymbol{\mu})$ (in **yellow**), the answers selected x_t could oscillate between one of the two correct answers marked by the **red** crosses.

Algorithm 1 Sticky-TaS [Degenne and Koolen, 2019]

Require: Total order over $\mathcal{X}$, exploration funct. $g(t)$

1: **Sampling Rule**
2: $C_t = \{\boldsymbol{\mu}' \in \mathcal{M} : D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \boldsymbol{\mu}') \leq \log(g(t))\}$
3: $\mathcal{X}_t = \bigcup_{\boldsymbol{\mu}' \in C_t} \mathcal{X}_F(\boldsymbol{\mu}')$
4: Pick $x_t \in \mathcal{X}_t$ according to the total order over $\mathcal{X}$
5: Compute $\boldsymbol{\omega}(t) \in \operatorname{argmax}_{\boldsymbol{\omega} \in \Delta_K} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{\omega}, \neg x_t)$
6: Let $\tilde{\boldsymbol{\omega}}(t)$ be the projection of $\boldsymbol{\omega}(t)$ onto $\Delta_K^{\epsilon_t} = \Delta_K \cap [\epsilon_t, 1]^K$
7: $A_t \in \operatorname{argmax}_{k \in [K]} \sum_{s=1}^t \omega_k(s) - N_k(t)$
8: **Stopping Rule**
9: Stop if $\beta_{t,\delta} < \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x)$
10: **Recommendation Rule**
11: $\hat{x}_{\tau_\delta} \in \operatorname{argmax}_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x)$

First, we first recall how Sticky-TaS works (pseudocode in Algorithm 1). During each round $t \in \mathbb{N}$, the algorithm defines a confidence region C_t around $\hat{\boldsymbol{\mu}}(t)$, using a suitable exploration function $g : \mathbb{N} \to \mathbb{R}$. Then, it computes a set of candidate answers $\mathcal{X}_t$ using models in C_t and selects an answer $x_t \in \mathcal{X}_t$ according to a pre-specified total order over $\mathcal{X}$. Once this is done, it computes the empirical oracle weights for answer x_t and it applies the C-Tracking [Garivier and Kaufmann, 2016] sampling rule on the sequence $\{\tilde{\boldsymbol{\omega}}(s)\}_{s=1}^t$, where $\tilde{\boldsymbol{\omega}}(s)$ denotes the l_∞ projection of $\boldsymbol{\omega}(s)$ onto $\Delta_K^{\epsilon_t}$ and $\epsilon_t = (4(t+K^2))^{-1/2}$. The algorithm then stops using the condition $\beta_{t,\delta} < \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x)$ and returns any answer that attains the argmax in $D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x)$. Here, $\beta_{t,\delta}$ is a calibrated threshold [Kaufmann and Koolen, 2021] which ensures δ-correctness.

The upper hemicontinuity of $\boldsymbol{\mu} \rightrightarrows \mathcal{X}_F(\boldsymbol{\mu})$ (point *(iv)* of Lemma 1) ensures that $\mathcal{X}_t \subseteq \mathcal{B}_\epsilon(\mathcal{X}_F(\boldsymbol{\mu}))$ for any $\epsilon > 0$ and sufficiently large t (indeed, C_t will concentrate around $\boldsymbol{\mu}$ under a good event). This property holds regardless of whether $\mathcal{X}$ is finite or infinite. However, when $\mathcal{X}$ is finite, we can essentially identify $\mathcal{B}_\epsilon(\mathcal{X}_F(\boldsymbol{\mu}))$ with $\mathcal{X}_F(\boldsymbol{\mu})$ and, therefore, we can say that $\mathcal{X}_t = \mathcal{X}_F(\boldsymbol{\mu})$ for sufficiently large t. Thus, Sticky TaS sticks to a fixed $x \in \mathcal{X}_t$ thanks to the pre-specified total order of $\mathcal{X}$. However, when $\mathcal{X}$ is not finite, one can only ensure that $\mathcal{X}_t \subseteq \mathcal{B}_\epsilon(\mathcal{X}_F(\boldsymbol{\mu}))$. As a consequence, the algorithm can fail at *sticking* to a single answer as the total order over $\mathcal{X}$ might select answers that progressively disappear from $\mathcal{X}_t$ (thus breaking the optimality proof of Sticky-TaS). Furthermore, a total order over $\mathcal{X}$ might even select answers in $\mathcal{X}_t$ in a way that prevents any sort of converging behavior to a single answer in $\mathcal{X}_F(\boldsymbol{\mu})$. In Figure 1 we present an example of this problem and in the next section we show how to circumvent it.[7]

5 AN OPTIMAL FRAMEWORK

In order to address the upper hemicontinuity issue discussed above, the main idea is selecting a sequence of candidate answers $x_t \in \mathcal{X}$ such that, under a good event, x_t progressively *converges* to some answer $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$.[8] Indeed, we will show that this guarantees asymptotically optimal rates. The rest of this section is structured as follows. In Section 5.1, we introduce a general framework (Sticky Sequence TaS) that works with any *converging selection rule* (Definition 1), and we present its theoretical guarantees. In Section 5.2, we show how to implement converging selection rules.

5.1 Sticky Sequence Track-and-Stop

Sticky Sequence Track-and-Stop shares the same pseudocode of Sticky Track-and-Stop (*i.e.*, Algorithm 1) with one major difference. Specifically, rather than using a pre-specified total order over $\mathcal{X}$ to select an answer $x_t \in \mathcal{X}_t$ (Line 4), it uses a *convergent selection rule*, which is defined as follows.

[7] One might argue that, although x_t is not converging, the corresponding oracle weights $\boldsymbol{\omega}(t)$ could still lead to optimality. In Appendix F we show with a theoretical example and an experiment that this is not the case.

[8] One might be tempted to solve the convergence issue of x_t by blindly discretizing the answer space, and then applying Sticky-TaS on the resulting finite answers problem. In Appendix F, we argue that this naive method does not attain asymptotic optimality.

Definition 1 (Convergent selection rule)**.** *A selection rule is said to be* convergent *if the sequence* $\{x_t\}_{t\in\mathbb{N}}$ *it generates satisfies the following property: for every* $\epsilon > 0$*, there exists a time* $T_\epsilon \in \mathbb{N}$ *such that, for all* $t \geq T_\epsilon$*, under the good event* $\mathcal{E}(t) = \bigcap_{s\geq h(t)}^{t} \{\boldsymbol{\mu} \in C_s\}$*, there exists* $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu}) : \|x_s - \bar{x}\| \leq \epsilon,\ \forall s \geq h(t) = \lceil \sqrt{t} \rceil$.

Intuitively, a convergent selection rule guarantees that, under the good event, the sequence $\{x_t\}_t$ stays close to a correct answer $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$. This is not a property guaranteed by *any* selection rule. For example, the selection rule employed by Sticky-TaS is not a convergent one when the answer set $\mathcal{X}$ is infinite, but it is convergent when $\mathcal{X}$ is finite (we provide more details on this in Section 5.2). We discuss in Section 5.2 how to implement selection rules that satisfies Definition 1. First, we prove that Sticky Sequence Track-and-Stop is asymptotically optimal whenever $\{x_t\}_t$ satisfies Definition 1.

Theorem 3. *Sticky Sequence Track-and-Stop, equipped with a convergent selection rule, is* δ*-correct and asymptotically optimal,* i.e.*,* $\limsup_{\delta\to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} \leq T^*(\boldsymbol{\mu})$.

Proof Sketch of Theorem 3 First, we observe that the stopping and recommendation rules lead to a δ-correct algorithm for any sampling rule even when the answer space is infinite, *i.e.*, for any sampling rule $(A_t)_{t\geq 1}$ we have that $\mathbb{P}_{\boldsymbol{\mu}}(\hat{x}_{\tau_\delta} \notin \mathcal{X}^\star(\boldsymbol{\mu})) \leq \delta$ for all $\boldsymbol{\mu} \in \mathcal{M}$ (Lemma 28). This result follows from standard concentration arguments. We now discuss how to prove asymptotic optimality. The proof, as usual in the literature, proceeds by analyzing the behavior of the algorithm under the sequence of events $\{\mathcal{E}(t)\}_t$ defined above. Specifically, we show that (i) $\sum_{t=0}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}(t)^c)$ is finite and (ii) there exists $T_0(\delta)$ such that $\mathcal{E}(t) \subseteq \{\tau_\delta \leq t\}$ for any $t \geq T_0(\delta)$, where $T_0(\delta)$ is such that $T_0(\delta)/\log(1/\delta) \to T^*(\boldsymbol{\mu})$ for $\delta \to 0$. Indeed, whenever these two conditions hold, one can prove asymptotic optimality with standard arguments. While (i) is well-known to be finite (see Degenne and Koolen [2019]), the crucial part is proving (ii). To this end, we recall that the algorithm stops as soon as it holds $\max_{x\in\mathcal{X}_F(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) \geq \beta_{t,\delta}$. Now, suppose that under the good event $\mathcal{E}(t)$ it holds:

$$\max_{x\in\mathcal{X}_F(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) \gtrsim tD(\boldsymbol{\mu}). \tag{2}$$

Let $T_0(\delta)$ be the first $t \in \mathbb{N}$ such that $tD(\boldsymbol{\mu}) \gtrsim \beta_{t,\delta}$ is satisfied. Then, for $t \geq T_0(\delta)$, we have $\mathcal{E}_t \subseteq \{\tau_\delta \leq t\}$. Furthermore, it is also possible to show that $T_0(\delta)/\log(1/\delta) \to T^*(\boldsymbol{\mu})$ for $\delta \to 0$. Thus, it only remains to show that Equation (2) holds under the good event $\mathcal{E}(t)$, which is the key step in the analysis. To this end, we will use three main ingredients. First, under event $\mathcal{E}(t)$, $\hat{\boldsymbol{\mu}}(t) \approx \boldsymbol{\mu}$ and $\boldsymbol{\mu}'(s) \approx \boldsymbol{\mu}$ for all $s \geq h(t)$. Second, C-tracking guarantees $\boldsymbol{N}(t) \approx \sum_{s=1}^{t} \boldsymbol{\omega}(s)$ (Lemma 41). Finally, and most importantly, x_s converges to some $\bar{x} \in \mathcal{X}^\star(\boldsymbol{\mu})$. These three key properties, along with the continuity results from Lemma 1, allow us to prove that:

$$\begin{aligned}\max_{x\in\mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) &\gtrsim D\left(\boldsymbol{\mu}, \sum_{s=1}^{t} \boldsymbol{\omega}(s), \neg\bar{x}\right) \\ &\geq \sum_{s=h(t)}^{t} D(\boldsymbol{\mu}, \boldsymbol{\omega}(s), \neg\bar{x}) \\ &\gtrsim \sum_{s=h(t)}^{t} D(\boldsymbol{\mu}'(s), \boldsymbol{\omega}(s), \neg x_s)\end{aligned}$$

where, (i) in the first step, we used $\boldsymbol{N}(t) \approx \sum_{s=1}^{t} \boldsymbol{\omega}(s)$, $\hat{\boldsymbol{\mu}}(t) \approx \boldsymbol{\mu}$, and $\bar{x} \in \mathcal{X}^\star(\boldsymbol{\mu})$ (ii), in the second one, we moved the inf inside the summation and we used the fact that $D \geq 0$ to get rid of the first $h(t)$ steps, and (iii) in the last one the convergence of x_s to $\bar{x}$ and $\boldsymbol{\mu}'(s) \approx \boldsymbol{\mu}$ for all s such that $s \geq h(t)$. Finally, we note that $D(\boldsymbol{\mu}'(s), \boldsymbol{\omega}(s), \neg x_s) = D(\boldsymbol{\mu}(s)')$ and thus that $D(\boldsymbol{\mu}'(s), \boldsymbol{\omega}(s), \neg x_s) \gtrsim D(\boldsymbol{\mu})$, thus concluding the proof.

We conclude by highlighting two important properties.

Generalization of existing algorithms It is important to highlight that Sticky Sequence Track-and-Stop generalizes both TaS and Sticky-TaS. Precisely, if x_t is chosen as a point in $\mathcal{X}_F(\hat{\boldsymbol{\mu}}(t))$, we obtain the TaS algorithm. While if x_t is selected according to a pre-specified total order over $\mathcal{X}$, then we obtain Sticky-TaS. In other words, both TaS and Sticky-TaS are selecting a sequence of answers x_t and collecting data to (eventually) identify x_t as a correct answer. However, as we discuss in the next section, the main point is that these selection rules can fail to generate a converging sequence, and therefore may fail to guarantee optimality in general settings with infinite answers.

Importance of convergence The proof sketch above critically relies on the fact that, for all $s \geq h(t)$, the points x_s remain close to a *fixed* element $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$. If this was not the case, *e.g.*, there exists $\bar{x}_s \in \mathcal{X}_F(\boldsymbol{\mu})$ such that $\bar{x}_s \approx x_s$ (this is always the case for both TaS and Sticky-TaS), then we would not be able to achieve the same result, as in step three we are "forced" to select a single answer within $\mathcal{X}_F(\boldsymbol{\mu})$ for *all* $s \geq h(t)$.[9]

[9] In Theorem 4 (Appendix E.4), we provide guarantees for the case in which the sequence of answers is not a converging one. Given the previous remark, we note that Theorem 4 provides theoretical guarantees on the performance

5.2 Algorithms for Converging Sequences

In this section, we discuss how to develop selection rules for the candidate answer x_t that ensure that $\{x_t\}_t$ is a converging sequence. Since, under the good event, we have $\mathcal{X}_t \subseteq \mathcal{B}_\epsilon(\mathcal{X}_F(\boldsymbol{\mu}))$, the problem reduces to the following: given a sequence of sets $\{\mathcal{X}_t\}_t$ such that $\mathcal{X}_t \to \mathcal{X}_F(\boldsymbol{\mu})$, how can we select a sequence of points $x_t \in \mathcal{X}_t$ such that $x_t \to \bar{x}$ for some $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$? We now discuss several solutions to this problem depending on the different topological properties of $\mathcal{X}$ and $\mathcal{X}_F(\boldsymbol{\mu})$. These different cases highlight (i) the challenges in building algorithms that satisfy Definition 1 and (ii) existing cases where existing methods, such as TaS and Sticky-TaS are asymptotically optimal. We remark that, in the last scenario that we cover at the end of this section, we consider the general setup where $\mathcal{X}$ is any subset of $\mathbb{R}^d$, hence providing a general procedure for building a convergent sequence.

When $\mathcal{X}_F(\mu)$ is single-valued First, when $\boldsymbol{\mu} \rightrightarrows \mathcal{X}_F(\boldsymbol{\mu})$ is single-valued for all $\boldsymbol{\mu} \in \Theta^K$, one can easily pick any answer within $\mathcal{X}_t$ to obtain a converging sequence. The underlying reason is that whenever $\boldsymbol{\mu} \rightrightarrows \mathcal{X}_F(\boldsymbol{\mu})$ is single-valued, then $\boldsymbol{\mu} \rightrightarrows \mathcal{X}_F(\boldsymbol{\mu})$ can be seen as a continuous function of $\boldsymbol{\mu}$.[10] Then, since, under the good event, $\mathcal{X}_t$ progressively converges to $\mathcal{X}_F(\boldsymbol{\mu})$, we have that picking any element $x_t \in \mathcal{X}_t$ leads to a converging sequence (Lemma 33). As a consequence, Theorem 3 implies that both TaS and Sticky-TaS are asymptotically optimal whenever $|\mathcal{X}_F(\boldsymbol{\mu})|$ is unique. The optimality of TaS for multiple-answer problems where $\mathcal{X}_F(\boldsymbol{\mu})$ is single-valued and $|\mathcal{X}|$ is finite was proved also in Degenne and Koolen [2019]. Theorem 3 extends the result to infinite-answer problems.

When $\mathcal{X} \subset \mathbb{R}$ Next, we study the case where $\mathcal{X} \subset \mathbb{R}$ (*e.g.*, this is the case when the goal is to estimate the optimal arm up an $\epsilon > 0$ accuracy). In this setting, selecting any $x_t \in \mathcal{X}_t$ clearly does not lead to a converging sequence. Consider, for instance, the case $x_t \in \mathcal{X}_F(\hat{\boldsymbol{\mu}}(t))$ (*i.e.*, the way in which TaS selects candidate answers) and $|\mathcal{X}_F(\boldsymbol{\mu})| = 2$. Since, the map $\mathcal{X}_F(\boldsymbol{\mu})$ is only upper hemicontinuous, it might be the case that for models $\boldsymbol{\mu}'$ arbitrarily close to $\boldsymbol{\mu}$, $\mathcal{X}_F(\boldsymbol{\mu}')$ only contains one of these two answers. As a consequence, since $x_t \in \mathcal{X}_F(\hat{\boldsymbol{\mu}}(t))$ and $\hat{\boldsymbol{\mu}}(t) \approx \boldsymbol{\mu}$, one might select different answers at each step t, which implies that $\{x_t\}_t$ does not converge.[11] However, a converging sequence can be obtained by picking $x_t \in \operatorname{argmin}_{x \in \mathcal{X}_t} x$ (alternatively, we could also take the max). It is easy to see that, whenever $\mathcal{X}_t \subseteq \mathcal{B}_\epsilon(\mathcal{X}_F(\boldsymbol{\mu}))$, then $|x_t - \operatorname{argmin}_{x \in \mathcal{X}_F(\boldsymbol{\mu})} x| \leq \epsilon$ (Lemma 34). In this sense, there exists a total order over $\mathcal{X}$ such that the resulting sequence is a converging one. As a consequence, from Theorem 3, we have that Sticky-TaS is optimal whenever $\mathcal{X} \subset \mathbb{R}$. However, as we will highlight in the next paragraph, this simple fix fails in dimension 2 and higher.

When $|\mathcal{X}_F(\boldsymbol{\mu})|$ is finite Now, consider the case where $|\mathcal{X}_F(\boldsymbol{\mu})|$ is finite (but $\mathcal{X}$ is possibly infinite). For the same arguments that we presented above, TaS fails at being optimal in this case (*e.g.*, if $|\mathcal{X}_F(\boldsymbol{\mu})| = 2$, but for points around $\boldsymbol{\mu}$ only one of these two answers is correct). Nevertheless, we now argue that even Sticky-TaS fails at generating a convergent sequence, *i.e.*, we cannot simply use a total order as above for picking answers. Consider $\mathcal{X} \subseteq \mathbb{R}^2$ and $\mathcal{X}_F(\boldsymbol{\mu}) = \{x_1, x_2\}$. Then, from the upper hemicontinuity of $\mathcal{X}_F(\boldsymbol{\mu})$, we have that, for sufficiently large t, on $\mathcal{E}(t)$, $\mathcal{X}_t \subseteq \mathcal{B}_\epsilon(x_1) \cup \mathcal{B}_\epsilon(x_2)$ and $\mathcal{B}_\epsilon(x_1) \cap \mathcal{B}_\epsilon(x_2) = \emptyset$. Then, selecting x_t using, *e.g.*, the lexicographic order over $\mathbb{R}^2$ might lead to oscillating behaviors between points in $\mathcal{B}_\epsilon(x_1)$ and $\mathcal{B}_\epsilon(x_2)$.[12] To fix this issue, we can resort to the following selection rule: $x_t \in \operatorname{argmin}_{x \in \mathcal{X}_t} \|x - x_{t-1}\|_\infty$. We prove in Lemma 35 that this leads to a converging sequence. Indeed, since for sufficiently large t, $\mathcal{X}_t \in \bigcup_{x \in \mathcal{X}_F(\boldsymbol{\mu})} \mathcal{B}_\epsilon(x)$ and $\mathcal{B}_\epsilon(x_1) \cap \mathcal{B}_\epsilon(x_2) = \emptyset$, by staying close to the previously selected point, we know that the algorithms will select points that are always within $\mathcal{B}_\epsilon(\bar{x})$ for some $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$. Interestingly, unlike in the previous cases, this procedure does not rely on exact knowledge of the points to which the algorithm will converge.

All the other cases Finally, we consider the most general case in which $\mathcal{X} \subset \mathbb{R}^d$. Here, one might be tempted to select again $x_t \in \operatorname{argmin}_{x \in \mathcal{X}_t} \|x - x_{t-1}\|_\infty$ in the hope of having again a convergent sequence. Nevertheless, suppose that $\mathcal{X}_F(\boldsymbol{\mu})$ is, *e.g.*, the boundary of the unitary ball centered in some $x \in \mathcal{X}$. Then, for sufficiently large t, we only have that $\mathcal{X}_t \subseteq \mathcal{B}_\epsilon(\mathcal{X}_F(\boldsymbol{\mu}))$. As a result, x_t might "wander" indefinitely around $\mathcal{X}_F(\boldsymbol{\mu})$ thus preventing convergence. To solve this issue, we propose an algorithm that progressively discretizes the answer space $\mathcal{X}$.[13] The key idea it to combine a *progressive discretization* of $\mathcal{X}$, using balls with a vanishing radius ρ_t over time, with a mechanism that incorporates the *history* of previously selected points. During each iteration, the al-

of TaS and Sticky-TaS for all cases where they fail to generate a converging sequence.

[10] Indeed, every upper hemicontinuous and single-valued correspondence is a continuous function.

[11] This was already spotted in Degenne and Koolen [2019].

[12] Consider again Figure 1, and suppose that the only answers in $\mathcal{X}_F(\boldsymbol{\mu})$ are the red crosses.

[13] For space constraints, its pseudocode can be found in the appendix (Algorithm 2; Appendix E.3.4); here, we outline only the general idea.

gorithm constructs and maintains a set $\mathcal{H}_t$ which is composed of tuples $(\bar{x}_s, \rho_s)_{s=1}^t$. Specifically, each element $(\bar{x}_s, \rho_s)$ represents a region $\mathcal{B}_{\rho_s}(\bar{x}_s)$ in the answer space in which the algorithm is "conducting the search" of an answer within $\mathcal{X}_F(\boldsymbol{\mu})$. More precisely, for all $s \in [1, t]$, the elements in $\mathcal{H}_t$ are such that $\mathcal{B}_{\rho_s}(\bar{x}_s) \cap \mathcal{B}_{\rho_{s-1}}(\bar{x}_{s-1}) \cap \mathcal{X}_t \neq \emptyset$, and $\bar{x}_s$ is the center of a ball radius ρ_s that belongs to a *uniquely* identified finite cover of $\mathcal{X}$, which we denote by $\mathcal{P}_s$. Since $\rho_s < \rho_{s-1}$, this implies that, over time, the regions $\mathcal{B}_{\rho_s}(\bar{x}_s)$ become progressively smaller. Given this setup, the algorithm simply selects x_t to be any element within $\mathcal{B}_{\rho_t}(\bar{x}_t) \cap \mathcal{X}_t$. To guide the search toward previously selected points, the set $\mathcal{H}_t$ is recursively constructed at each iteration as follows. The algorithm first selects the point $(\bar{x}, \bar{\rho})$ in $\mathcal{H}_{t-1}$ with smaller radius and for which it holds that $\mathcal{B}_{\rho_s}(\bar{x}_s) \cap \mathcal{X}_t \neq \emptyset$ for all $(\bar{x}_s, \rho_s) \in \mathcal{H}_{t-1}$ such that $\rho_s < \bar{\rho}$ (indeed, as iteration progresses, previously selected balls might not contain any candidate answers within the updated $\mathcal{X}_t$). Then, the algorithm constructs $\mathcal{H}_t$ starting from $(\bar{x}, \bar{\rho})$ in order to satisfy $\mathcal{B}_{\rho_s}(\bar{x}_s) \cap \mathcal{B}_{\rho_{s-1}}(\bar{x}_{s-1}) \cap \mathcal{X}_t \neq \emptyset$. Such a procedure can be proven to generate a converging sequence (Lemma 38). The reason why this happens is that, due to the upper hemicontinuity of $\mathcal{X}_F(\boldsymbol{\mu})$, $\mathcal{X}_t \in \mathcal{B}_\epsilon(\mathcal{X}_F(\boldsymbol{\mu}))$ for sufficiently large t. This implies that, for sufficiently small ϵ and sufficiently large t, balls of radius $\mathcal{B}_\epsilon(\bar{x})$ that belong to $\mathcal{P}_\epsilon$ and that do not intersect $\mathcal{X}_F(\boldsymbol{\mu})$ will not belong to $\mathcal{H}_t$. As a consequence, the algorithm has found a region of radius ϵ that contains an element of $\mathcal{X}_F(\boldsymbol{\mu})$ and this region will remain in $\mathcal{H}_t$ under the good event.

6 FUTURE WORK

Our study paves the way for several future research avenues. For example, one might investigate what happens outside of regular pure exploration problems. Are these problems learnable or is Assumption 3 necessary for finite sample complexity? Furthermore, while our algorithm is statistically optimal, we observe that it is not computationally efficient. This limitation is somewhat expected, given that even Sticky-TaS [Degenne and Koolen, 2019] (which deals with a narrower class of problems) has analogous limitations. A promising direction for future work is to investigate whether efficient algorithms can be developed for *specific* classes of infinite-answer pure exploration problems (*e.g.*, the regression of a continuous function $f(\boldsymbol{\mu})$). Indeed, despite Sticky-TaS being inefficient, there exist algorithms that are both efficient and optimal for the ε-best arm identification problem [Jourdan et al., 2023]. Another possible approach is to focus on improving computational complexity by relaxing asymptotic optimality guarantees and instead targeting β-optimal algorithms, which are usually more efficient [Russo, 2016, Qin et al., 2017, Jourdan et al., 2022].

Acknowledgements

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Work supported by the Cariplo CRYPTONOMEX grant and by an ERC grant (Project 101165466 — PLA-STEER).

References

Jean-Pierre Aubin. *Set-Valued Analysis*. Birkhäuser Boston, 1999.

Claude Berge. Espaces topologiques: Fonctions multivoques. *Dunod, Paris*, 1959.

Olivier Cappé, Aurélien Garivier, Odalric-Ambrym Maillard, Rémi Munos, and Gilles Stoltz. Kullback-leibler upper confidence bounds for optimal sequential allocation. *The Annals of Statistics*, pages 1516–1541, 2013.

Emil Carlsson, Debabrota Basu, Fredrik Johansson, and Devdatt Dubhashi. Pure exploration in bandits with linear constraints. In *International Conference on Artificial Intelligence and Statistics*, pages 334–342. PMLR, 2024.

Vikas Deep, Achal Bassamboo, and Sandeep Kumar Juneja. Asymptotically optimal and computationally efficient average treatment effect estimation in a/b testing. In *Forty-first International Conference on Machine Learning*, 2024.

Rémy Degenne and Wouter M Koolen. Pure exploration with multiple correct answers. *Advances in Neural Information Processing Systems*, 32, 2019.

Rémy Degenne, Wouter M Koolen, and Pierre Ménard. Non-asymptotic pure exploration by solving games. *Advances in Neural Information Processing Systems*, 32, 2019.

Rémy Degenne, Pierre Ménard, Xuedong Shang, and Michal Valko. Gamification of pure exploration for linear bandits. In *International Conference on Machine Learning*, pages 2432–2442. PMLR, 2020a.

Rémy Degenne, Han Shao, and Wouter Koolen. Structure adaptive algorithms for stochastic bandits. In *International Conference on Machine Learning*, pages 2443–2452. PMLR, 2020b.

Eyal Even-Dar, Shie Mannor, and Yishay Mansour. Pac bounds for multi-armed bandit and markov decision processes. In *Computational Learning Theory: 15th Annual Conference on Computational Learning Theory, COLT 2002 Sydney, Australia, July 8–10, 2002 Proceedings 15*, pages 255–270. Springer, 2002.

Eugene A Feinberg, Pavlo O Kasyanov, and Mark Voorneveld. Berge's maximum theorem for noncompact image sets. *Journal of Mathematical Analysis and Applications*, 413(2):1040–1046, 2014.

Aurélien Garivier and Emilie Kaufmann. Optimal best arm identification with fixed confidence. In *Conference on Learning Theory*, pages 998–1027. PMLR, 2016.

Marc Jourdan and Rémy Degenne. Choosing answers in epsilon-best-answer identification for linear bandits. In *International Conference on Machine Learning*, pages 10384–10430. PMLR, 2022.

Marc Jourdan, Mojmír Mutnỳ, Johannes Kirschner, and Andreas Krause. Efficient pure exploration for combinatorial bandits with semi-bandit feedback. In *Algorithmic Learning Theory*, pages 805–849. PMLR, 2021.

Marc Jourdan, Rémy Degenne, Dorian Baudry, Rianne de Heide, and Emilie Kaufmann. Top two algorithms revisited. *Advances in Neural Information Processing Systems*, 35:26791–26803, 2022.

Marc Jourdan, Rémy Degenne, and Emilie Kaufmann. An ε-best-arm identification algorithm for fixed-confidence and beyond. *Advances in Neural Information Processing Systems*, 36:16578–16649, 2023.

Sandeep Juneja and Subhashini Krishnasamy. Sample complexity of partition identification using multi-armed bandits. In *Conference on Learning Theory*, pages 1824–1852. PMLR, 2019.

Emilie Kaufmann and Wouter M Koolen. Mixture martingales revisited with applications to sequential tests and confidence intervals. *Journal of Machine Learning Research*, 22(246):1–44, 2021.

Tomáš Kocák and Aurélien Garivier. Best arm identification in spectral bandits. *arXiv preprint arXiv:2005.09841*, 2020.

Tomás Kocák and Aurélien Garivier. Epsilon best arm identification in spectral bandits. In *IJCAI*, pages 2636–2642, 2021.

Joseph Lazzaro and Ciara Pike-Burke. Fixed-confidence multiple change point identification under bandit feedback. *arXiv preprint arXiv:2507.08994*, 2025.

Erich L Lehmann and George Casella. *Theory of point estimation*. Springer Science & Business Media, 2006.

Arnab Maiti. Open problem: Optimal instance-dependent sample complexity for finding nash equilibrium in two player zero-sum matrix games. *Proceedings of Machine Learning Research vol*, 291:1–5, 2025.

Arnab Maiti, Kevin Jamieson, and Lillian Ratliff. Instance-dependent sample complexity bounds for zero-sum matrix games. In *International Conference on Artificial Intelligence and Statistics*, pages 9429–9469. PMLR, 2023.

Arnab Maiti, Ross Boczar, Kevin Jamieson, and Lillian Ratliff. Near-optimal pure exploration in matrix games: A generalization of stochastic bandits & dueling bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 2602–2610. PMLR, 2024.

Pierre Ménard. Gradient ascent for active exploration in bandit problems. *arXiv preprint arXiv:1905.08165*, 2019.

Vrettos Moulos. Optimal best markovian arm identification with fixed confidence. *Advances in Neural Information Processing Systems*, 32, 2019.

Takayuki Osogami, Junya Honda, and Junpei Komiyama. Optimal estimation of the best mean in multi-armed bandits. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Riccardo Poiani, Rémy Degenne, Emilie Kaufmann, Alberto Maria Metelli, and Marcello Restelli. Optimal multi-fidelity best-arm identification. *Advances in Neural Information Processing Systems*, 37:121882–121927, 2024.

Riccardo Poiani, Marc Jourdan, Emilie Kaufmann, and Rémy Degenne. Best-arm identification in unimodal bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 2233–2241. PMLR, 2025.

Chao Qin, Diego Klabjan, and Daniel Russo. Improving the expected improvement algorithm. *Advances in Neural Information Processing Systems*, 30, 2017.

Yoan Russac, Christina Katsimerou, Dennis Bohle, Olivier Cappé, Aurélien Garivier, and Wouter M Koolen. A/b/n testing with control in the presence of subpopulations. *Advances in Neural Information Processing Systems*, 34:25100–25110, 2021.

Alessio Russo, Yichen Song, and Aldo Pacchiano. Pure exploration with feedback graphs. *arXiv preprint arXiv:2503.07824*, 2025.

Daniel Russo. Simple bayesian algorithms for best arm identification. In *Conference on learning theory*, pages 1417–1418. PMLR, 2016.

Adrienne Tuynman and Rémy Degenne. The batch complexity of bandit pure exploration. *arXiv preprint arXiv:2502.01425*, 2025.

Filippo Vannella, Alexandre Proutiere, and Jaeseong Jeong. Best arm identification in multi-agent multi-armed bandits. In *International Conference on Machine Learning*, pages 34875–34907. PMLR, 2023.

Po-An Wang, Ruo-Chun Tzeng, and Alexandre Proutiere. Fast pure exploration via frank-wolfe. *Advances in Neural Information Processing Systems*, 34:5810–5821, 2021.

Yichi Zhou, Jialian Li, and Jun Zhu. Identify the nash equilibrium in static games with random payoffs. In *International Conference on Machine Learning*, pages 4160–4169. PMLR, 2017.

Checklist

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. Yes. We formally described the problem in
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. We have analyzed important properties of the algorithm such as the sample complexity. Time and space complexity are not applicable as the proposed method are not applicable as they require to fix a specific pure exploration problem.
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. Yes, we provided anonymized source code for the code used in the experiments as supplementary material.
2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. Yes, assumptions are formally stated for each theoretical result.
   (b) Complete proofs of all theoretical results. Yes, we provide both complete proofs in the appendix. We also included proof sketches for the main theoretical results (some of these have been deferred to the appendix due to space constraints).
   (c) Clear explanations of any assumptions. Yes, all assumptions are clearly stated. Why they are needed and their interpretation is explained throughout the main text.
3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). Yes, we provided these details in the supplementary materials.
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). Yes, we report these details in the experiment section in the appendix.
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). Yes, we report these details in the experiment section in the appendix.
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). Yes, we report these details in the experiment section in the appendix.
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
   (a) Citations of the creator If your work uses existing assets. Not Applicable. Our work does not use existing assets.
   (b) The license information of the assets, if applicable. Not Applicable. Our work does not use existing assets.
   (c) New assets either in the supplemental material or as a URL, if applicable. Yes. We included the code used for the experiments in the supplemental material.
   (d) Information about consent from data providers/curators. Not Applicable.
   (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. Not Applicable.
5. If you used crowdsourcing or conducted research with human subjects, check if you include:
   (a) The full text of instructions given to participants and screenshots. Not Applicable. We did not used crowdsourcing and we did not conducted research with human subjects.
   (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. Not Applicable. We did not used crowdsourcing and we did not conducted research with human subjects.
   (c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. Not Applicable. We did not used crowdsourcing and we did not conducted research with human subjects.

Supplementary Material: Pure Exploration with Infinite Answers

Appendix Content

A Examples of Regular Pure Exploration Problems **13**

A.1 Regression of Continuous Functions . . . 13

A.2 Finite Set of Answers . . . 13

A.3 Composition of Learning Problems . . . 14

A.4 ϵ-Nash Equilibrium in Two-player Zero-sum Games . . . 15

B On Regular Pure Exploration Problems **17**

B.1 On Assumption 2 . . . 17

B.2 On Assumption 3 . . . 18

C Proof of the Lower Bound **23**

C.1 Proof Sketch of Theorem 2 . . . 23

C.2 Proof of Theorem 2 . . . 24

D Continuity Results **27**

E Algorithm Analysis **29**

E.1 Correctness and Expected Stopping Time . . . 29

E.2 Sticky-Sequence Track-and-Stop . . . 30

E.3 Algorithms for Converging Sequences . . . 34

E.4 What happens when $\{x_t\}$ is not converging . . . 40

F On simple solutions **43**

F.1 Why discretizing the answer space does not work . . . 44

F.2 Empirical Failure of Sticky Track-and-Stop . . . 44

G Helper Lemmas **46**

G.1 Tracking . . . 46

G.2 Stopping time . . . 46

H Mathematical Background **47**

H.1 Set Theory . . . 47

H.2 Canonical Exponential Family . . . 47

H.3 Results on the infimum . . . 49

H.4 Correspondences . . . 49

A Examples of Regular Pure Exploration Problems

In this section, we present several general classes of problems for which Assumptions 1 to 3 hold. Specifically, the first class of problems is that of regression of continuous functions (Appendix A.1), the second one encompasses all finite-answers problems (Appendix A.2), and the third one consists of combinations of problems that meet Assumptions 1 to 3 (Appendix A.3). The last class allows one to combine regression of continuous functions together with finite-answers problems.

Some results of this section rely on Theorem 1, which is proved in Appendix B.

A.1 Regression of Continuous Functions

Consider a continuous function $f : \Theta^K \to \mathcal{X}$ and let $\epsilon > 0$. Suppose that both Θ^K and $\mathcal{X}$ are compact Euclidean sets. The goal of the learner is to compute an ϵ-accurate estimate of $f(\boldsymbol{\mu})$ with high-probability, namely, $\mathbb{P}_{\boldsymbol{\mu}}(\|f(\boldsymbol{\mu}) - x_{\tau_\delta}\|_\infty > \epsilon) \leq \delta$ for all $\boldsymbol{\mu} \in \mathcal{M}$. Thus, $\mathcal{X}^\star(\boldsymbol{\mu}) = \{x \in \mathcal{X} : \|f(\boldsymbol{\mu}) - x\|_\infty \leq \epsilon\}$.

Lemma 2. *Assumption 1 holds, namely $\mathcal{X}$ is compact and $\boldsymbol{\mu} \rightrightarrows \mathcal{X}^\star(\boldsymbol{\mu})$ is compact-valued.*

Proof. The set $\mathcal{X}$ is compact by assumption, and $\mathcal{X}^\star(\boldsymbol{\mu}) = \{x \in \mathcal{X} : \|f(\boldsymbol{\mu}) - x\|_\infty \leq \epsilon\}$ is compact as well. □

We now continue by proving that Assumption 2 holds.

Lemma 3. *Assumption 2 holds. Namely, for all $\boldsymbol{\mu} \in \mathcal{M}$, there exists $\bar{x} \in \mathcal{X}^\star(\boldsymbol{\mu})$ such that $\boldsymbol{\mu} \notin \mathrm{cl}(\neg\bar{x})$*

Proof. Let $\mathcal{Q}_x = \{\boldsymbol{\mu} \in \mathcal{M} : \|f(\boldsymbol{\mu}) - x\|_\infty < \epsilon\}$. We first prove that $\mathcal{Q}_x \cap \mathrm{cl}(\neg x) = \emptyset$. We recall that a closure of a generic set $\mathcal{S}$ contains all the points in $\mathcal{S}$ together with all the limit points. Now, for all $\boldsymbol{\mu} \in \neg x$ we have $\|f(\boldsymbol{\mu}) - x\| > \epsilon$, hence $\mathcal{Q}_x \cap \neg x = \emptyset$. It remains to prove that no limit point of $\neg x$ belongs to $\mathcal{Q}_x$. Suppose, by contradiction, that $\boldsymbol{\mu}$ is a limit point of $\mathrm{cl}(\neg x)$ and that $\boldsymbol{\mu} \in \mathcal{Q}_x$. Then, then there exists a sequence $\boldsymbol{\mu}_n$ such that $\boldsymbol{\mu}_n \in \neg x$ and $\boldsymbol{\mu}_n \to \boldsymbol{\mu}$. However, by continuity of f we would have $\|f(\boldsymbol{\mu}_n) - x\|_\infty \to \|f(\boldsymbol{\mu}) - x\|_\infty < \epsilon$, thus showing that $\boldsymbol{\mu}_n$ is not in $\neg x$ for sufficiently large n.

Now, for any $\boldsymbol{\mu} \in \mathcal{M}$, let $\bar{x} = f(\boldsymbol{\mu})$. Then, $\bar{x} \in \mathcal{X}^\star(\boldsymbol{\mu})$ since $\|f(\boldsymbol{\mu}) - \bar{x}\|_\infty = 0 \leq \epsilon$. Furthermore, $\boldsymbol{\mu} \notin \mathrm{cl}(\neg\bar{x})$ (since $\boldsymbol{\mu} \in \mathcal{Q}_{\bar{x}}$ and $\mathcal{Q}_{\bar{x}} \cap \mathrm{cl}(\neg\bar{x}) = \emptyset$), which concludes the proof. □

Finally, we continue by proving Assumption 3. We prove that by simply showing that $\mathcal{X}^\star(\boldsymbol{\mu})$ is a continuous correspondence.

Lemma 4. *The correspondence $\boldsymbol{\mu} \rightrightarrows \mathcal{X}^\star(\boldsymbol{\mu})$ is continuous over Θ^K.*

Proof. To prove this statement, we use a sufficient condition for the continuity of the composition of a correspondence and a continuous function. In particular, we can use Theorem 7. Indeed the correspondence $\mathcal{X}^\star(\boldsymbol{\mu})$ can be seen as $\mathcal{X}^\star(\boldsymbol{\mu}) = \cup_{u \in \mathcal{B}_\epsilon(0)}\{f(\boldsymbol{\mu}) + u\}$. Thus, $\mathcal{X}^\star : \Theta^K \rightrightarrows \mathcal{X}$ is continuous. □

Lemma 5. *Assumption 3 holds, namely for all $\epsilon > 0$ sufficiently small, there exists $\rho > 0$ such that for all $\boldsymbol{\mu} \in \Theta^K$, $\boldsymbol{\omega} \in \Delta_K$, $x \in \mathcal{X}$, it holds that $\neg\mathcal{B}_\rho(x) \neq \emptyset$ and $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg\mathcal{B}_\rho(x)) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x) \leq \epsilon$.*

Proof. Apply Lemma 4 and Theorem 1. □

A.2 Finite Set of Answers

We show that Assumptions 1 to 3 hold in the setting where the answer space is finite Degenne and Koolen [2019]. Specifically, we consider $\mathcal{M} \subseteq \Theta^K$ and let $|\mathcal{X}| < +\infty$, and we consider an arbitrary correspondence $\mathcal{X}^\star : \Theta^K \rightrightarrows \mathcal{X}$ that models the set of correct answers for each bandit model. Without loss of generality, we only assume that Assumption 2 holds. Otherwise, one cannot obtain finite sample-complexity results (see Appendix B.1). That being said, we now prove that Assumption 1 and Assumption 3 hold.

Lemma 6. *Assumption 1 holds; namely $\mathcal{X}$ is compact and $\mathcal{X}^*(\boldsymbol{\mu})$ is compact-valued.*

Proof. $\mathcal{X}$ is compact since it is finite. Therefore, $\mathcal{X}^\star(\boldsymbol{\mu})$ is compact as well since $\mathcal{X}^\star(\boldsymbol{\mu}) \subseteq \mathcal{X}$. □

Lemma 7. *Assumption 3 holds; namely, for all $\epsilon > 0$ sufficiently small, there exists $\rho > 0$ such that for all $\boldsymbol{\mu} \in \Theta^K$, $\boldsymbol{\omega} \in \Delta_K$, $x \in \mathcal{X}$, it holds that $\neg\mathcal{B}_\rho(x) \neq \emptyset$ and $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg\mathcal{B}_\rho(x)) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x) \leq \epsilon$.*

Proof. We first show that there exists $\rho > 0$ such that $\neg\mathcal{B}_\rho(x) \neq \emptyset$. Since $\mathcal{X}$ is finite, for any $x \in \mathcal{X}$, there exists $\rho_x > 0$ such that $\mathcal{B}_{\rho_x}(x) = \{x\}$. Then, taking $\rho = \min_{x\in\mathcal{X}} \rho_x$, we obtain that $\neg\mathcal{B}_\rho(x) = \neg x$, $\forall x \in \mathcal{X}$. Furthermore, $\neg x \neq \emptyset$ holds by design otherwise one could simply always return such x.

Now, taking $\rho = \min_{x\in\mathcal{X}} \rho_x$, for all $\epsilon > 0$ it follows that:

$$D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg\mathcal{B}_\rho(x)) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x) = 0 \leq \epsilon,$$

thus concluding the proof. □

A.3 Composition of Learning Problems

We now consider the composition of learning problems, where, for each bandit $\boldsymbol{\mu} \in \Theta^K$ the set of correct answer $\mathcal{X}^\star(\boldsymbol{\mu})$ is the product correspondence of two correspondences $\mathcal{X}_1^\star(\boldsymbol{\mu}) \times \mathcal{X}_2^\star(\boldsymbol{\mu})$, where $\mathcal{X}_1^\star(\boldsymbol{\mu}) \subseteq \mathcal{X}_1$, $\mathcal{X}_2^\star(\boldsymbol{\mu}) \subseteq \mathcal{X}_2$, and $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$. Specifically, we consider

$$\mathcal{X}^\star(\boldsymbol{\mu}) = \{x = (x_1, x_2) \in \mathcal{X} : x_1 \in \mathcal{X}_1^\star(\boldsymbol{\mu}) \text{ and } x_2 \in \mathcal{X}_2^\star(\boldsymbol{\mu})\} = \mathcal{X}_1^\star(\boldsymbol{\mu}) \times \mathcal{X}_2^\star(\boldsymbol{\mu}).$$

For the sake of the example, consider the case where we want to estimate the $\min_{k\in[K]} \mu_k$ up to an ϵ-factor, together with the problem of finding an arm k that satisfies $\mu_k > \operatorname{argmax}_{j\in[K]} \mu_k - \epsilon$. This problem is a joint combination of a regression problem together with a multiple (but finite) correct answers one. As we show in this section, this sort of problem is regular.

Indeed, suppose that Assumption 1, Assumption 2 and Assumption 3 hold for the learning problems defined by $\mathcal{X}_1^\star$ and $\mathcal{X}_2^\star$ independently. Furthermore, let $\neg_1, \neg_2$ be the alternative set correspondences related to $\mathcal{X}_1^\star$ and $\mathcal{X}_2^\star$, respectively. Similarly, let $\neg$ be the correspondence related to the product correspondence $\mathcal{X}^\star$. Suppose that $\neg_1 x_1 \neq \emptyset$ for all $x_1 \in \mathcal{X}_1$ and $\neg_2 x_2 \neq \emptyset$ for all $x_2 \in \mathcal{X}_2$, so that $\neg x \neq \emptyset$ as well.

We now show that the learning problem defined by $\mathcal{X}^\star$ inherits Assumption 1, Assumption 2 and Assumption 3 directly from $\mathcal{X}_1^\star$ and $\mathcal{X}_2^\star$.

Lemma 8. *Assumption 1 holds; namely, $\mathcal{X}$ is compact and $\mathcal{X}^\star$ is compact-valued.*

Proof. Given two compact sets A, B, $A \times B$ is compact. Thus, since $\mathcal{X}_1, \mathcal{X}_2$ are compact, $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$ is compact as well. Furthermore, since $\mathcal{X}^\star(\boldsymbol{\mu}) = \mathcal{X}_1^\star(\boldsymbol{\mu}) \times \mathcal{X}_2^\star(\boldsymbol{\mu})$, and $\mathcal{X}_1^\star(\boldsymbol{\mu}), \mathcal{X}_2^\star(\boldsymbol{\mu})$ are compact, then $\mathcal{X}^\star(\boldsymbol{\mu})$ is compact as well. □

Lemma 9. *Assumption 2 holds; namely for all $\boldsymbol{\mu} \in \mathcal{M}$, there exists $\bar{x} \in \mathcal{X}^\star(\boldsymbol{\mu})$ such that $\boldsymbol{\mu} \notin \operatorname{cl}(\neg\bar{x})$.*

Proof. For all $\boldsymbol{\mu} \in \mathcal{M}$, there exists $\bar{x}_1 \in \mathcal{X}_1^\star(\boldsymbol{\mu})$ such that $\boldsymbol{\mu} \notin \operatorname{cl}(\neg_1\bar{x}_1)$, and, moreover, there exists $\bar{x}_2 \in \mathcal{X}_2^\star(\boldsymbol{\mu})$ such that $\boldsymbol{\mu} \notin \operatorname{cl}(\neg_2\bar{x}_2)$.

Now, consider $\bar{x} = (\bar{x}_1, \bar{x}_2) \in \mathcal{X}^\star(\boldsymbol{\mu})$. We verify that $\bar{x} \notin \operatorname{cl}(\neg\bar{x})$. Specifically,

$$\begin{aligned}
\operatorname{cl}(\neg\bar{x}) &= \operatorname{cl}\left(\{\boldsymbol{\lambda} \in \mathcal{M} : \bar{x} \notin \mathcal{X}^\star(\boldsymbol{\mu})\}\right) \\
&= \operatorname{cl}\left(\{\boldsymbol{\lambda} \in \mathcal{M} : \bar{x}_1 \notin \mathcal{X}_1^\star(\boldsymbol{\mu}) \text{ or } \bar{x}_2 \notin \mathcal{X}_2^\star(\boldsymbol{\mu})\}\right) \\
&= \operatorname{cl}\left(\{\boldsymbol{\lambda} \in \mathcal{M} : \bar{x}_1 \notin \mathcal{X}_1^\star(\boldsymbol{\mu})\} \cup \{\boldsymbol{\lambda} \in \mathcal{M} : \bar{x}_2 \notin \mathcal{X}_2^\star(\boldsymbol{\mu})\}\right) \\
&= \operatorname{cl}(\neg_1\bar{x}_1) \cup \operatorname{cl}(\neg_2\bar{x}_2),
\end{aligned}$$

where in the last step we used that, for any two sets A, B, $\operatorname{cl}(A \cup B) = \operatorname{cl}(A) \cup \operatorname{cl}(B)$.

Now, since $\boldsymbol{\mu} \notin \operatorname{cl}(\neg_1\bar{x}_1)$ and $\boldsymbol{\mu} \notin \operatorname{cl}(\neg_2\bar{x}_2)$, it follows that $\boldsymbol{\mu} \notin \operatorname{cl}(\neg\bar{x})$, thus concluding the proof. □

Lemma 10. *Assumption 3 holds; namely, for all $\epsilon > 0$ sufficiently small, there exists $\rho > 0$ such that $\neg\mathcal{B}_\rho(x) \neq \emptyset$ and for all $\boldsymbol{\mu} \in \Theta^K, \boldsymbol{\omega} \in \Delta_K, x \in \mathcal{X}$, it holds that $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg\mathcal{B}_\rho(x)) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x) \leq \epsilon$.*

Proof. We first show that there exists ρ such that $\neg\mathcal{B}_\rho(x) \neq \emptyset$, $\forall x \in \mathcal{X}$. Given $x = (x_1, x_2)$, we have:

$$\begin{aligned}
\neg\mathcal{B}_\rho(x) &= \{\boldsymbol{\lambda} \in \mathcal{M} : \forall \bar{x} \in \mathcal{B}_\rho(x), \bar{x} \notin \mathcal{X}^\star(\boldsymbol{\lambda})\} \\
&= \{\boldsymbol{\lambda} \in \mathcal{M} : \forall \bar{x} = (\bar{x}_1, \bar{x}_2) \in \mathcal{B}_\rho(x), \bar{x}_1 \notin \mathcal{X}_1^\star(\boldsymbol{\lambda}) \text{ or } \bar{x}_2 \notin \mathcal{X}_2^\star(\boldsymbol{\lambda})\} \\
&= \{\boldsymbol{\lambda} \in \mathcal{M} : \forall \bar{x}_1 \in \mathcal{B}_\rho(x_1), \bar{x}_1 \notin \mathcal{X}_1^\star(\boldsymbol{\lambda})\} \cup \{\boldsymbol{\lambda} \in \mathcal{M} : \forall \bar{x}_2 \in \mathcal{B}_\rho(x_2), \bar{x}_2 \notin \mathcal{X}_2^\star(\boldsymbol{\lambda})\} \\
&= (\neg_1\mathcal{B}_\rho(x_1)) \cup (\neg_2\mathcal{B}_\rho(x_2)) \,.
\end{aligned}$$

Now, since Assumption 3 holds for $\mathcal{X}_1^\star, \mathcal{X}_2^\star$, there exists $\bar{\rho}_1, \bar{\rho}_2 > 0$ such that $\neg_1\mathcal{B}_\rho(x_1), \neg_2\mathcal{B}_\rho(x_2)$ are both non empty for all $\rho \leq \min\{\bar{\rho}_1, \bar{\rho}_2\}$.[14]

Now, we continue by analyzing the difference in the divergences. Proceeding as above, we have that:

$$\begin{aligned}
\neg x &= \neg_1 x_1 \cup \neg_2 x_2 \\
\neg\mathcal{B}_\rho(x) &= \neg_1\mathcal{B}_\rho(x_1) \cup \neg_2\mathcal{B}_\rho(x_2).
\end{aligned}$$

Now, let $\epsilon > 0$, and set $\rho = \min\{\rho_1, \rho_2\}$, where ρ_1 and ρ_2 are such that:

$$D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg_1\mathcal{B}_{\rho_1}(x_1)) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg_1 x_1) \leq \epsilon \quad \forall \boldsymbol{\mu} \in \Theta^K, \boldsymbol{\omega} \in \Delta_K, x_1 \in \mathcal{X}_1 \tag{3}$$

$$D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg_2\mathcal{B}_{\rho_2}(x_2)) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg_2 x_2) \leq \epsilon \quad \forall \boldsymbol{\mu} \in \Theta^K, \boldsymbol{\omega} \in \Delta_K, x_2 \in \mathcal{X}_2 \tag{4}$$

Then, let $\boldsymbol{\mu} \in \Theta^K$, $\boldsymbol{\omega} \in \Delta_K$, $x = (x_1, x_2) \in \mathcal{X}$. For readability, since $\boldsymbol{\mu}, \boldsymbol{\omega}$ are fixed, we omit them from the notation $D(\cdot,\cdot,\cdot)$ in the rest of this proof. Suppose without loss of generality that $D(\neg_1 x_1) \leq D(\neg_2 x_2)$. It holds that:

$$\begin{aligned}
D(\neg\mathcal{B}_\rho(x)) - D(\neg x) &= D(\neg_1\mathcal{B}_\rho(x_1) \cup \neg_2\mathcal{B}_\rho(x_2)) - D(\neg_1 x_1 \cup \neg_2 x_2) \\
&= D(\neg_1\mathcal{B}_\rho(x_1) \cup \neg_2\mathcal{B}_\rho(x_2)) - \min\{D(\neg_1 x_1), D(\neg_2 x_2)\} \\
&= D(\neg_1\mathcal{B}_\rho(x_1) \cup \neg_2\mathcal{B}_\rho(x_2)) - D(\neg_1 x_1) \\
&\leq D(\neg_1\mathcal{B}_\rho(x_1)) - D(\neg_1 x_1) \\
&\leq D(\neg_1\mathcal{B}_{\rho_1}(x_1)) - D(\neg_1 x_1) \\
&\leq \epsilon
\end{aligned}$$

where (i) in the first step we used the decomposition of $\neg$ into $\neg_1$ and $\neg_2$, (ii) in the second one, Lemma 46, (iii) in the third one $D(\neg_1 x_1) \leq D(\neg_2 x_2)$, (iv) in the forth one $\neg_1\mathcal{B}_\rho(x_1) \subseteq \neg_1\mathcal{B}_\rho(x_1) \cup \neg_2\mathcal{B}_\rho(x_2)$, and (v) in the last step Equation (3). □

A.4 ϵ-Nash Equilibrium in Two-player Zero-sum Games

We now consider the problem of learning an ϵ-Nash equilibrium point in a two player zero-sum stochastic matrix game [Zhou et al., 2017, Maiti et al., 2023, 2024, Maiti, 2025]. Specifically, we consider the following setting. We let $\Theta^K = \mathcal{M} = [-1, 1]^{n\times n}$ where $n \in \mathbb{N}$ denotes the number of actions for each player and we suppose that each distribution is *e.g.*, Gaussian.[15] Each point $\boldsymbol{\mu} \in \mathcal{M}$ denotes the mean of a stochastic payoff matrix whose distributions are in the underlying canonical exponential family. At each round, the learner controls both the min player and the max player by selecting an action $(I_t, J_t) \in n \times n$, and it receives as payoff $R_t \sim \nu_{i,j}$, where $\nu_{i,j}$ is the distribution with mean $\mu_{i,j}$. In the rest of this section, we slightly overload the notation and we denote by x the min player and by y the max player. Let $\mathcal{X} = \Delta_n \times \Delta_n$ be the answer space. Then, we define the agent's goal as follows:

$$\mathcal{X}^\star(\boldsymbol{\mu}) = \left\{(x, y) \in \Delta_n^2 : \max_{j\in[n]}(x^\top\boldsymbol{\mu})_j - \min_{i\in[n]}(\boldsymbol{\mu}y)_j \leq \epsilon\right\}, \tag{5}$$

where $\epsilon > 0$ is the desired accuracy parameter and $(\boldsymbol{\mu}y)_i = \sum_{j\in[n]} y_j\mu_{i,j}$ and $(x^\top\boldsymbol{\mu})_j = \sum_{i\in[n]} x_i\mu_{i,j}$.

We now show that this learning task is a regular pure exploration problem.

[14] Notice that one could take also the maximum of $\bar{\rho}_1, \bar{\rho}_2$, since $\neg\mathcal{B}_\rho(x)$ is the union of the two $\neg_1\mathcal{B}_\rho(x_1)$ and $\neg_2\mathcal{B}_\rho(x_2)$.

[15] The following argument can easily be extended to the $n \times m$ setting and to other bounded domains outside the $[-1, 1]$ interval provided that Θ is a closed interval strictly contained in an open one.

Lemma 11. *Assumption 1 holds; namely, $\mathcal{X}$ is compact and $\mathcal{X}^\star$ is compact valued.*

Proof. Recall that $\mathcal{X} = \Delta_n \times \Delta_n$, which is clearly compact, and $\mathcal{X}^\star(\boldsymbol{\mu})$ is compact as well. □

Lemma 12. *Assumption 2 holds; namely, for all $\boldsymbol{\mu} \in \mathcal{M}$, there exists $(\bar{x}, \bar{y}) \in \mathcal{X}^\star(\boldsymbol{\mu})$ such that $\boldsymbol{\mu} \notin \mathrm{cl}(\neg(\bar{x}, \bar{y}))$.*

Proof. Fix $\boldsymbol{\mu} \in \mathcal{M}$ and consider $(x^\star, y^\star) \in \Delta_n^2$ such that $(x^\star, y^\star)$ is an exact Nash equilibrium for $\boldsymbol{\mu}$ (guaranteed to exists for each payoff matrix $\boldsymbol{\mu}$). At $(x^\star, y^\star)$ it holds that $(x^\star)^\top \boldsymbol{\mu} e_j \leq (x^\star)^\top \boldsymbol{\mu} y^\star \leq e_i^\top \boldsymbol{\mu} y^\star$. This, in turns, implies that, $(x^\star, y^\star) \in \mathcal{X}^\star(\boldsymbol{\mu})$. We now prove that $\boldsymbol{\mu} \notin \mathrm{cl}(\neg(x^\star, y^\star))$. Recall that:

$$\begin{aligned}\mathrm{cl}(\neg(x^\star, y^\star)) &= \mathrm{cl}\left(\{\boldsymbol{\lambda} \in \mathcal{M} : (x^\star, y^\star) \notin \mathcal{X}^\star(\boldsymbol{\lambda})\}\right) \\ &= \mathrm{cl}\left(\left\{\boldsymbol{\lambda} \in \mathcal{M} : \max_{j\in[n]}((x^\star)^\top \boldsymbol{\lambda})_j - \min_{i\in[n]}(\boldsymbol{\lambda} y^\star)_i > \epsilon\right\}\right),\end{aligned}$$

and proceed by contradiction. Assume that $\boldsymbol{\mu} \in \mathrm{cl}(\neg(x^\star, y^\star))$. Then, there must exist a sequence $\{\boldsymbol{\mu}_n\}_n \to \boldsymbol{\mu}$ such that $\boldsymbol{\mu}_n \neq \boldsymbol{\mu}$ and $\boldsymbol{\mu}_n \in \mathrm{cl}(\neg(x^\star, y^\star))$ $\forall n$. However, if $\boldsymbol{\mu}_n \to \boldsymbol{\mu}$, the continuity of $\boldsymbol{\lambda} \to \max_{j\in[n]}((x^\star)^\top \boldsymbol{\lambda})_j - \min_{i\in[n]}(\boldsymbol{\lambda} y^\star)_i$, would imply that

$$\max_{j\in[n]}((x^\star)^\top \boldsymbol{\mu}_n)_j - \min_{i\in[n]}(\boldsymbol{\mu}_n y^\star)_i \to \max_{j\in[n]}((x^\star)^\top \boldsymbol{\mu})_j - \min_{i\in[n]}(\boldsymbol{\mu} y^\star)_i = 0.$$

However, for all n, $\max_{j\in[n]}((x^\star)^\top \boldsymbol{\mu}_n)_j - \min_{i\in[n]}(\boldsymbol{\mu}_n y^\star)_i > \epsilon > 0$, thus leading to a contradiction. □

Lemma 13. *Assumption 3 holds; namely, for all $\epsilon > 0$ sufficiently small, there exists $\rho > 0$ such that $\neg\mathcal{B}_\rho((x, y)) \neq \emptyset$ and for all $\boldsymbol{\mu} \in \Theta^K, \boldsymbol{\omega} \in \Delta_K, (x, y) \in \mathcal{X}$, it holds that $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg\mathcal{B}_\rho((x, y))) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg(x, y)) \leq \epsilon$.*

Proof. We this result using Theorem 1. Thus, we only need to prove the continuity of the correspondence $\boldsymbol{\mu} \rightrightarrows \mathcal{X}^\star(\boldsymbol{\mu})$. We start by proving that Equation (5) is upper hemicontinuous.[16] To do that, we note that $\mathcal{X}$ is compact and we use Theorem 8, which states that $\mathcal{X}^\star(\boldsymbol{\mu})$ is upper hemicontinuous if its graph is closed. Thus, we need to prove that

$$\mathrm{Gr}(\mathcal{X}^\star) = \left\{(\boldsymbol{\mu}, (x, y)) \in [-1, 1]^{n\times n} \times \Delta_n^2 : \max_{j\in[n]}(x^\top \boldsymbol{\mu})_j - \min_{i\in[n]}(\boldsymbol{\mu} y)_j \leq \epsilon\right\}$$

is closed, *i.e.*, that it contains all its limit points. Let $(\bar{\boldsymbol{\mu}}, (\bar{x}, \bar{y}))$ be any limit point of $\mathrm{Gr}(\mathcal{X}^\star)$. Then, there exists $\{(\boldsymbol{\mu}_n, (x_n, y_n))\}_n$ such that (i) $(\boldsymbol{\mu}_n, (x_n, y_n)) \in \mathrm{Gr}(\mathcal{X}^\star)$, (ii) $(\boldsymbol{\mu}_n, (x_n, y_n)) \neq (\bar{\boldsymbol{\mu}}, (\bar{x}, \bar{y}))$ and (iii) $(\boldsymbol{\mu}_n, (x_n, y_n)) \to (\bar{\boldsymbol{\mu}}, (\bar{x}, \bar{y}))$. Let us introduce some notation for brevity. Specifically, we define the following function $(\boldsymbol{\mu}, (x, y)) \to F(\boldsymbol{\mu}, (x, y))$

$$F(\boldsymbol{\mu}, (x, y)) = \max_{j\in[n]}(x^\top \boldsymbol{\mu})_j - \min_{i\in[n]}(\boldsymbol{\mu} y)_j.$$

Then, let $\bar{F} = F(\bar{\boldsymbol{\mu}}, (\bar{x}, \bar{y}))$ and $F_n = (\boldsymbol{\mu}_n, (x_n, y_n))$. Now, suppose that $\bar{F} = \bar{\epsilon} > \epsilon$, *i.e.*, that $(\bar{\boldsymbol{\mu}}, (\bar{x}, \bar{y})) \notin \mathrm{Gr}(\mathcal{X}^\star(\boldsymbol{\mu})))$. For any n it holds that:

$$\bar{F} = \bar{F} - F_n + F_n \leq |\bar{F} - F_n| + \epsilon,$$

where we used that $F_n \leq \epsilon$ since $(\boldsymbol{\mu}_n, (x_n, y_n)) \in \mathrm{Gr}(\mathcal{X}^\star)$. Now, from the continuity of $(\boldsymbol{\mu}, x, y) \to F_n$, we have that $F_n \to \bar{F}$ as $n \to +\infty$. One can now pick n such that $|\bar{F} - F_n| \leq (\bar{\epsilon} - \epsilon)/2$, thus leading to $\bar{F} \leq (\bar{\epsilon} - \epsilon)/2 + \epsilon < \bar{\epsilon}$, hence contradicting the fact that $\bar{F} = \bar{\epsilon} > \epsilon$. Thus, the graph is closed and Equation (5) is upper hemicontinuous.

We now continue by proving the lower hemicontinuity. We start with a preliminary consideration. Specifically, we observe that $\mathcal{X}^\star(\boldsymbol{\mu})$ is a polytope with non-empty interior for all $\boldsymbol{\mu} \in [-1, 1]^{n\times n}$. Indeed, we have that:

$$\begin{aligned}\mathcal{X}^\star(\boldsymbol{\mu}) &= \left\{(x, y) \in \Delta_n^2 : \max_{j\in[n]}(x^\top \boldsymbol{\mu})_j - \min_{i\in[n]}(\boldsymbol{\mu} y)_j \leq \epsilon\right\} \\ &= \left\{(x, y) \in \Delta_n^2 : (x^\top \boldsymbol{\mu})_j - (\boldsymbol{\mu} y)_i \leq \epsilon, \forall i, j \in [n]^2\right\}.\end{aligned}$$

[16] Before reading the proof, the we reader might find useful to read Appendix B.

Now, for a fixed $\boldsymbol{\mu}$, each constraint is linear in x, y and Δ_n^2 is in itself a polytope. We now prove that the interior is non-empty. Here, it is sufficient to consider an exact Nash equilibrium point $(x^\star, y^\star)$; at $(x^\star, y^\star)$ it holds that $(x^\star)^\top \boldsymbol{\mu} e_j \leq (x^\star)^\top \boldsymbol{\mu} y^\star \leq e_i^\top \boldsymbol{\mu} y^\star$, and thus $F(\boldsymbol{\mu}, (x^\star, y^\star)) = 0 < \epsilon$ and the interior is non-empty. Furthermore, for a fixed $\boldsymbol{\mu}$, $F(\boldsymbol{\mu}, (x, y))$ is convex in (x, y). Indeed, we can rewrite F as follows:

$$F(\boldsymbol{\mu}, (x, y)) = \max_{i,j \in [n]^2} \left((x^\top \boldsymbol{\mu})_j - (\boldsymbol{\mu} y)_i \right),$$

which is a maximum over a finite set of linear function, which is known to be convex.

To prove the lower hemicontinuity, we need to show that, for any $\boldsymbol{\mu} \in [-1, 1]^{n \times n}$ and any open set $\mathcal{V} \in \Delta_n^2$ such that $\mathcal{X}^\star(\boldsymbol{\mu}) \cap \mathcal{V} \neq \emptyset$, there exists a neighbourhood U of $\boldsymbol{\mu}$ such that $\forall \boldsymbol{\mu}' \in U$, $\mathcal{X}^\star(\boldsymbol{\mu}') \cap \mathcal{V} \neq \emptyset$. Fix a neighbourhood U of $\boldsymbol{\mu}$ or radius κ (we will pick an appropriate κ later in the proof), *i.e.*, $\|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \kappa$ for all $\boldsymbol{\mu}' \in U$. Then, for any $(x, y) \in \Delta_n^2$ it holds that:

$$\begin{aligned} F(\boldsymbol{\mu}', (x, y)) &= \max_{j \in [n]} (x^\top \boldsymbol{\mu}')_j - \min_{i \in [n]} (\boldsymbol{\mu}' y)_j \\ &= \max_{j \in [n]} (x^\top (\boldsymbol{\mu}' - \boldsymbol{\mu} + \boldsymbol{\mu}))_j - \min_{i \in [n]} ((\boldsymbol{\mu}' - \boldsymbol{\mu} + \boldsymbol{\mu}) y)_j \\ &\leq F(\boldsymbol{\mu}, (x, y)) + 2\kappa. \end{aligned}$$

Thus, a necessary condition for having $(x, y) \in \mathcal{X}^\star(\boldsymbol{\mu}')$ is that $F(\boldsymbol{\mu}, (x, y)) + 2\kappa \leq \epsilon$.

Now, fix any point $(x, y) \in \mathcal{X}^\star(\boldsymbol{\mu}) \cap \mathcal{V}$ and let $(x^\star, y^\star)$ be an exact Nash equilibrium such that $F(\boldsymbol{\mu}, (x^\star, y^\star)) = 0$. Define (x_α, y_α) as a convex combination between (x, y) and $(x^\star, y^\star)$ with parameter $\alpha \in [0, 1]$. Then, since $\mathcal{X}^\star(\boldsymbol{\mu})$ is a polytope, $(x_\alpha, y_\alpha) \in \mathcal{X}^\star(\boldsymbol{\mu})$ and, furthermore, since $(x, y) \in \mathcal{V}$ and $\mathcal{V}$ is open, there exists $\alpha \in (0, 1)$ such that $(x_\alpha, y_\alpha) \in \mathcal{V}$ as well. Now, consider $F(\boldsymbol{\mu}, (x_\alpha, y_\alpha))$. By the convexity of F, it holds that:

$$F(\boldsymbol{\mu}, (x_\alpha, y_\alpha)) \leq (1 - \alpha) F(\boldsymbol{\mu}, (x, y)) + F(\boldsymbol{\mu}, (x^\star, y^\star)) = (1 - \alpha) F(\boldsymbol{\mu}, (x, y)) \leq (1 - \alpha)\epsilon.$$

Thus, we have that $(x_\alpha, y_\alpha) \in \mathcal{V} \cap \mathcal{X}^\star(\boldsymbol{\mu}')$ for all $\boldsymbol{\mu}' \in U$ if $(1 - \alpha)\epsilon + 2\kappa \leq \epsilon$. Choosing $\kappa = \frac{\epsilon\alpha}{4}$ concludes the proof. □

Remark on the problem of learning ϵ-Nash equilibrium In the proof of Lemma 13, for the lower hemicontinuity of $\mathcal{X}^\star(\boldsymbol{\mu})$, we exploited the fact that the correspondence defines a polytope for each payoff matrix $\boldsymbol{\mu}$. This is true only if we define ϵ-Nash strategies in which the sum of the regrets is less than ϵ, which has the advantage of canceling the value of the game $x^\top \boldsymbol{\mu} y$ (which is bilinear) from the definition. The same trick would not work for general-sum games, in which the sum of the regrets remains bilinear. We leave as future work to handle such cases.

B On Regular Pure Exploration Problems

In this section, we show that Assumption 2 is necessary for finite sample complexity results (Appendix B.1), and we show that Assumption 3 holds for continuous correspondences (Appendix B.2).

B.1 On Assumption 2

In this section, we discuss Assumption 2. As we anticipated in the main text, Assumption 2 is necessary for obtaining finite sample-complexity results. Indeed, even for problems with finitely many possible answers, the failure of Assumption 2 leads to infinite lower bounds on the sample complexity. This is a direct consequence of Theorem 1 in Degenne and Koolen [2019].

Indeed, suppose that $\mathcal{X}$ is finite and that Assumption 2 does not hold. Then, there exists $\bar{\boldsymbol{\mu}} \in \mathcal{M}$ such that $\bar{\boldsymbol{\mu}} \in \mathrm{cl}(\neg x)$ for all $x \in \mathcal{X}^\star(\bar{\boldsymbol{\mu}})$. However, from Theorem 1 in Degenne and Koolen [2019], this would imply that:

$$\begin{aligned} &\liminf_{\delta \to 0} \frac{\mathbb{E}_{\bar{\boldsymbol{\mu}}}[\tau_\delta]}{\log(1/\delta)} \geq T^*(\bar{\boldsymbol{\mu}}) = D(\bar{\boldsymbol{\mu}})^{-1}, \\ &D(\bar{\boldsymbol{\mu}}) = \max_{x \in \mathcal{X}^\star(\bar{\boldsymbol{\mu}})} \max_{\boldsymbol{\omega} \in \Delta_K} \inf_{\boldsymbol{\lambda} \in \neg x} \sum_{k \in [K]} \omega_k d(\bar{\mu}_k, \lambda_k). \end{aligned}$$

However, since $\bar{\boldsymbol{\mu}} \in \mathrm{cl}(\neg x)$ for all $x \in \mathcal{X}^\star(\bar{\boldsymbol{\mu}})$, this would lead to $D(\bar{\boldsymbol{\mu}}) = 0$, and hence $T^*(\bar{\boldsymbol{\mu}}) = +\infty$, thus leading to a lower bound with infinite sample complexity.

The above claims are related to the following result, showing that $D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}) > 0$ holds if and only if $\boldsymbol{\mu} \notin \mathrm{cl}(\neg\widetilde{\mathcal{X}})$.

Lemma 14 (Strictly positive divergence)**.** *Let $\boldsymbol{\mu} \in \mathcal{M}$ and $\widetilde{\mathcal{X}} \subseteq \mathcal{X}$. Then, $D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}) > 0$ holds if and only if $\boldsymbol{\mu} \notin \mathrm{cl}(\neg\widetilde{\mathcal{X}})$.*

Proof. We prove the first direction by contradiction. Suppose that $D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}) > 0$, that is $\sup_{\boldsymbol{\omega}\in\Delta_K} \inf_{\boldsymbol{\lambda}\in\neg\widetilde{\mathcal{X}}} \sum_{k\in[K]} \omega_k d(\mu_k, \lambda_k) > 0$ and suppose by contradiction that $\boldsymbol{\mu} \in \mathrm{cl}(\neg\widetilde{\mathcal{X}})$. Then we can take a sequence $\boldsymbol{\lambda}^j$ such that $\boldsymbol{\lambda}^j \to \boldsymbol{\mu}$ and, since $\sum_{k\in[K]} \omega_k d(\mu_k, \cdot)$ is continuous for all $\boldsymbol{\omega} \in \Delta_K$ and $\boldsymbol{\mu} \in \mathcal{M}$, we obtain $\sum_{k\in[K]} \omega_k d(\mu_k, \lambda_k^j) \to \sum_{k\in[K]} \omega_k d(\mu_k, \mu_k) = 0$, which shows that $D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}) = 0$.

Now, we prove the second direction. Suppose that $\boldsymbol{\mu} \notin \mathrm{cl}(\neg\widetilde{\mathcal{X}})$. Then, there exists $\epsilon > 0$, such that, for all $\boldsymbol{\lambda} \in \mathrm{cl}(\neg\widetilde{\mathcal{X}})$, $\|\boldsymbol{\mu} - \boldsymbol{\lambda}\|_\infty \geq \epsilon$. Thus, for all $\boldsymbol{\lambda} \in \mathrm{cl}(\neg\widetilde{\mathcal{X}})$, there exists $k \in [K]$ such that $|\mu_k - \lambda_k| > 0$, and consequently, $d(\mu_k, \lambda_k) > 0$. It follows that:

$$D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}) \geq \frac{1}{K} \inf_{\boldsymbol{\lambda}\in\neg\widetilde{\mathcal{X}}} \sum_{k\in[K]} d(\mu_k, \lambda_k) > 0,$$

which concludes the proof. □

B.2 On Assumption 3

In the following, we show that Assumption 3 holds under very mild conditions. Specifically, it is sufficient to assume that (i) $\boldsymbol{\mu} \rightrightarrows \mathcal{X}^\star(\boldsymbol{\mu})$ is a continuous and compact-valued correspondence, and (ii) that $\mathcal{M}$ and $\mathcal{X}$ are compact sets.

The structure of this section is organized as follows. First, in Appendix B.2.1, we prove that there exists $\rho > 0$ such that $\forall x \in \mathcal{X}$, $\neg\mathcal{B}_\rho(x) \neq \emptyset$. As one can note, this is the first statement in Assumption 3. Secondly, in Appendix B.2.2, we provide some intermediate and helper results that are used to prove the second part of Assumption 3. Finally, in Appendix B.2.3, we prove the second part of Assumption 3, that is the fact that for all $\epsilon > 0$, there exists ρ such that $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg\mathcal{B}_\rho(x)) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x) \leq \epsilon$ holds uniformly across models, weights and answers.

B.2.1 (Extended) Alternative Models are non-empty, that is $\neg\mathcal{B}_\rho(x) \neq \emptyset$

We start by showing that, under the assumptions above there exists $\rho > 0$ such that $\neg\mathcal{B}_\rho(x) \neq \emptyset$ for all $x \in \mathcal{X}$.

Lemma 15. *There exists $\rho > 0$, such that, for all $x \in \mathcal{X}$, $\neg\mathcal{B}_\rho(x) \neq \emptyset$.*

Proof. Let $\bar{x} \in \mathcal{X}$. By definition, $\neg\mathcal{B}_\rho(\bar{x})$ is empty if, for all $\boldsymbol{\mu} \in \mathcal{M}$, there exists $x \in \mathcal{B}_\rho(\bar{x})$ such that $x \in \mathcal{X}^\star(\boldsymbol{\mu})$. In the following, we show that this cannot happen for an appropriate value of ρ. This value is derived in the following constructive way.

Consider the following function $g : \mathcal{X} \to \mathbb{R}$ defined as follows:

$$g(x) = \sup_{\boldsymbol{\lambda}\in\mathcal{M}} \inf_{\tilde{x}\in\mathcal{X}^\star(\boldsymbol{\lambda})} \|x - \tilde{x}\|_\infty.$$

Then, by Theorem 5, since $\lambda \rightrightarrows \mathcal{X}^\star(\boldsymbol{\lambda})$ is continuous and compact-valued, and since $\|x - \tilde{x}\|_\infty$ is continuous, we have that $\inf_{\tilde{x}\in\mathcal{X}^\star(\boldsymbol{\lambda})} \|x - \tilde{x}\|_\infty$ is continuous as well. Furthermore, the inf is attained since it is an infimum of a continuous function over a compact domain. Furthermore, the supremum is also attained since the inf is continuous and $\mathcal{M}$ is compact. Finally, $g(x)$ is continuous. This follows from the fact $\mathcal{M}$ is constant (and hence a continuous correspondence of $\boldsymbol{\lambda}$) and compact-valued. Then, by further applying Theorem 5, we have proved that $x \to g(x)$ is continuous.

Now, consider $\boldsymbol{\lambda}_x \in \mathrm{argmax}_{\boldsymbol{\lambda}\in\mathcal{M}} \inf_{\tilde{x}\in\mathcal{X}^\star(\boldsymbol{\lambda})} \|x - \tilde{x}\|_\infty$. We note that $\min_{\tilde{x}\in\mathcal{X}^\star(\boldsymbol{\lambda}_x)} \|x - \tilde{x}\|_\infty > 0$. Indeed, that minimum is 0 if and only if $x \in \mathcal{X}^\star(\boldsymbol{\lambda}_x)$. Nevertheless, for all $x \in \mathcal{X}$, there exists $\bar{\boldsymbol{\lambda}} \in \neg x$, and, therefore,

$\min_{\tilde{x}\in\mathcal{X}^\star(\bar{\boldsymbol{\lambda}})} \|x-\tilde{x}\|_\infty > 0$. Thus, we have that:

$$\min_{\tilde{x}\in\mathcal{X}^\star(\boldsymbol{\lambda}_x)} \|x-\tilde{x}\|_\infty \geq \min_{\tilde{x}\in\mathcal{X}^\star(\bar{\boldsymbol{\lambda}})} \|x-\tilde{x}\|_\infty > 0.$$

Now, consider $\eta = \inf_{x\in\mathcal{X}} g(x)$. Since $g(x)$ is continuous, and since $\mathcal{M}$ is compact, then the inf is attained, and, as a consequence $\eta > 0$.

The proof then follows by picking $\rho = \frac{\eta}{2}$. Indeed, suppose that there exists $\bar{x} \in \mathcal{X}$ such that $\neg\mathcal{B}_\rho(\bar{x}) = \emptyset$. Then, for all $\boldsymbol{\mu} \in \mathcal{M}$, there exists $x_{\boldsymbol{\mu}} \in \mathcal{X} : \|\bar{x} - x_{\boldsymbol{\mu}}\|_\infty \leq \rho$ and $x_{\boldsymbol{\mu}} \in \mathcal{X}^\star(\boldsymbol{\mu})$. However, this would imply that:

$$\eta = \min_{x\in\mathcal{X}} g(x) \leq \min_{\tilde{x}\in\mathcal{X}^\star(\boldsymbol{\lambda}_{\bar{x}})} \|\bar{x}-\tilde{x}\|_\infty \leq \|\bar{x} - x_{\boldsymbol{\lambda}_{\bar{x}}}\|_\infty \leq \frac{\eta}{2},$$

since $\eta > 0$ this leads to a contradiction, thus concluding the proof. □

B.2.2 Preliminary Results

In order to continue, we need some intermediate results. Before proceeding, we recall that, given a set $\mathcal{S} \subseteq \widetilde{\mathcal{S}}$, $s \in \mathcal{S}$ is either a limit point of $\mathcal{S}$ or an isolated point. A point is isolated if $s \in \mathcal{S}$ and there exists a neighborhood $\mathcal{U}$ of s such that $\mathcal{U} \cap \mathcal{S} = \{s\}$. On the other hand, a point $s \in \widetilde{S}$ is a limit point of $\mathcal{S}$ if every neighbourhood $\mathcal{U}$ of s contains at least one point of $\mathcal{S}$ different from s itself. When dealing with metric spaces, this is equivalent to saying that there exists a sequence of points in $\mathcal{S} \setminus \{s\}$ whose limit is s.

That being said, we now consider the correspondence $x \rightrightarrows \mathrm{cl}(\neg x)$.

Lemma 16. *The correspondence $x \rightrightarrows \mathrm{cl}(\neg x)$ is lower hemicontinuous and compact-valued over $\mathcal{X}$.*

Proof. First, we note that $\mathrm{cl}(\neg x)$ is compact-valued. The set is trivially bounded and closed. Thus, it is compact by the Heine-Borel theorem.

We continue by proving that it is lower hemicontinuous. Consider an open set $\mathcal{V}$ and $x \in \mathcal{X}$ such that $\mathcal{V} \cap \mathrm{cl}(\neg x) \neq \emptyset$. Then, since $\mathcal{V}$ is open and $\mathrm{cl}(\neg x)$ is compact, we have that there exists $\boldsymbol{\lambda} \in \neg x$ such that $\boldsymbol{\lambda} \in \mathcal{V} \cap \mathrm{cl}(\neg x)$. To prove this, we proceed by contradiction. Assume that $\neg x \cap \mathcal{V} = \emptyset$. Then, since there exists $\boldsymbol{\lambda} \in \mathcal{V} \cap \mathrm{cl}(\neg x)$, then it must hold that $\boldsymbol{\lambda}$ is a limit point of $\mathrm{cl}(\neg x)$, i.e., there exists $\{\boldsymbol{\lambda}_n\}_{n\geq 1}$ such that $\boldsymbol{\lambda}_n \neq \boldsymbol{\lambda}$, $\boldsymbol{\lambda}_n \in \neg x$ and $\boldsymbol{\lambda}_n \to \boldsymbol{\lambda}$. Therefore, for all $\alpha > 0$, there exists $\bar{\boldsymbol{\lambda}} \in \neg x$ and $\|\boldsymbol{\lambda} - \bar{\boldsymbol{\lambda}}\|_\infty \leq \alpha$. But, then, since $\mathcal{V}$ is open and $\boldsymbol{\lambda} \in \mathcal{V}$, we can take α sufficiently small so that $\bar{\boldsymbol{\lambda}} \in \mathcal{V}$ as well.

Now, consider $\boldsymbol{\lambda} \in \neg x$ and $\boldsymbol{\lambda} \in \mathcal{V} \cap \neg x$. In the following, we will prove that there exists a neighborhood $\mathcal{B}_\kappa(x)$ of x such that $\boldsymbol{\lambda} \in \neg x'$ for all $x' \in \mathcal{B}_\kappa(x)$. Define $\kappa \in \mathbb{R}$ as follows:

$$\kappa = \frac{1}{2} \min_{\tilde{x}\in\mathcal{X}^\star(\boldsymbol{\lambda})} \|x-\tilde{x}\|_\infty > 0,$$

where in the inequality step we used $x \notin \mathcal{X}^\star(\boldsymbol{\lambda})$. Now, consider $x' \in \mathcal{X}$ such that $\|x - x'\| \leq \kappa$. Then, it holds that $\boldsymbol{\lambda} \in \neg x'$. Indeed, suppose it is false. Then:

$$\|x-x'\|_\infty \leq \kappa = \frac{1}{2} \min_{\tilde{x}\in\mathcal{X}^\star(\boldsymbol{\lambda})} \|x-\tilde{x}\|_\infty \leq \frac{1}{2}\|x-x'\|_\infty < \|x-x'\|_\infty,$$

thus leading to a contradiction. Thus, $\boldsymbol{\lambda} \in \neg x'$ and, by definition $\boldsymbol{\lambda} \in \mathcal{V}$. Thus, $\mathrm{cl}(\neg x') \cap \mathcal{V} \neq \emptyset$ and the correspondence is lower hemicontinuous. □

In the following lemma, we begin by characterizing isolated points of the graph of the correspondence $\mathrm{cl}\neg$. Before that, we introduce some notation. Consider a correspondence $\phi : \mathbb{X} \rightrightarrows \mathbb{Y}$. Then, for any $Z \subseteq \mathbb{X}$, let

$$\mathrm{Gr}_Z(\phi) = \{(x,y) \in Z \times \mathbb{Y} : y \in \phi(x)\}.$$

In the following, we will provide some characterization of $\mathrm{Gr}_{\mathcal{X}}(\mathrm{cl}\neg)$

Lemma 17. *Let $\mathrm{Gr}_{\mathcal{X}}(\mathrm{cl}\neg) = \{(x,\boldsymbol{\lambda}) \in \mathcal{X} \times \mathcal{M} : \boldsymbol{\lambda} \in \mathrm{cl}(\neg x)\}$ and $(x,\boldsymbol{\lambda}) \in \mathrm{Gr}_{\mathcal{X}}(\mathrm{cl}\neg)$ be an isolated point of $\mathrm{Gr}_{\mathcal{X}}(\mathrm{cl}\neg)$. Then $\boldsymbol{\lambda}$ is an isolated point of $\mathrm{cl}(\neg x)$.*

Proof. Since $(x, \boldsymbol{\lambda}) \in \mathrm{Gr}_{\mathcal{X}}(\mathrm{cl}\neg)$ is an isolated point of $\mathrm{Gr}_{\mathcal{X}}(\mathrm{cl}\neg)$, we have that, $\exists \bar{\kappa} > 0$ such that, for all $\kappa \in (0, \bar{\kappa}]$, $\mathcal{B}_\kappa((x, \boldsymbol{\lambda})) \cap \mathrm{Gr}_{\mathcal{X}}(\mathrm{cl}\neg) = (x, \boldsymbol{\lambda})$.

Thus, for all $\kappa \in (0, \bar{\kappa}]$ and all $\bar{\boldsymbol{\lambda}}$ such that $\|\bar{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \kappa$, $\bar{\boldsymbol{\lambda}} \neq \boldsymbol{\lambda}$, we have that $(x, \bar{\boldsymbol{\lambda}}) \notin \mathrm{Gr}_{\mathcal{X}}(\mathrm{cl}\neg)$. It follows that $\bar{\boldsymbol{\lambda}} \notin \mathrm{cl}(\neg x)$ as well. Thus, $\boldsymbol{\lambda}$ is an isolated point of $\mathrm{cl}(\neg x)$, thus concluding the proof. □

Next, we characterize isolated points of $\mathrm{cl}(\neg x)$.

Lemma 18. *Let $x \in \mathcal{X}$ and let $\boldsymbol{\lambda} \in \mathrm{cl}(\neg x)$ be an isolated point. Then $\boldsymbol{\lambda} \in \neg x$.*

Proof. This follows trivially from the definition of isolated points. Indeed, by definition, $\mathrm{cl}(\neg x)$ is defined as the union of $\neg x$ together with all the limit points of $\neg x$. Since $\boldsymbol{\lambda}$ is an isolated point, it is not a limit point. Thus, $\boldsymbol{\lambda} \in \neg x$. □

B.2.3 Continuous Correspondences Imply Assumption 3

Given the preliminary results discussed so far, we now dive into proving the second part of Assumption 3, *i.e.*, the fact that, for all $\epsilon > 0$ there exists $\rho > 0$ such that $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg\mathcal{B}_\rho(x)) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x) \leq \epsilon$ holds uniformly across Θ, Δ_K and $\mathcal{X}$.

Proof Sketch The main idea behind the proof is showing that, for all $\eta > 0$, there exists $\bar{\rho}_\eta > 0$ such that, for all $x \in \mathcal{X}$, $\boldsymbol{\lambda} \in \mathrm{cl}(\neg x)$, $\rho \in (0, \bar{\rho}_\eta]$, there exists $\tilde{\boldsymbol{\lambda}} \in \neg\mathcal{B}_\rho(x)$ such that $\|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \eta$ (Lemma 22). Indeed, whenever this condition holds, Assumption 3 holds as well (Lemma 23). It is important to note that, for all $\eta > 0$, $\bar{\rho}_\eta$ needs to uniformly exists for all $x \in \mathcal{X}$ and $\boldsymbol{\lambda} \in \mathrm{cl}(\neg x)$. That being said, the proof of the existence of $\bar{\rho}_\eta$ is constructive, and requires the following intermediate steps (Lemma 19, Lemma 20 and Lemma 21). Specifically, we will show that it is sufficient to set $\bar{\rho}_\eta$ as follows:

$$\bar{\rho}_\eta = \frac{1}{2} \min_{x \in \mathcal{X}} \min_{\boldsymbol{\lambda} \in \mathrm{cl}(\neg x)} \max_{\tilde{\boldsymbol{\lambda}} \in \mathcal{M} : \|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \eta} \min_{\tilde{x} \in \mathcal{X}^\star(\tilde{\boldsymbol{\lambda}})} \|x - \tilde{x}\|_\infty.$$

In this sense, Lemma 19, Lemma 20 and Lemma 21 analyze this expression with a bottom-up approach whose main goal is showing that the value of the optimization problem is strictly greater than 0. Once this is done, Lemma 22 proves that for any $\boldsymbol{\lambda} \in \mathrm{cl}(\neg x)$, there exists $\tilde{\boldsymbol{\lambda}} \in \neg\mathcal{B}_\rho(x)$ such that $\|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \eta$ for all $\rho \leq \bar{\rho}_\eta$, which, in turn, will imply Assumption 3 via Lemma 23.

Now that the main steps of the proof have been highlighted, we present the following lemma that study the optimization problem:

$$g_\eta(x, \boldsymbol{\lambda}) = \sup_{\tilde{\boldsymbol{\lambda}} \in \mathcal{M} : \|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \eta} \inf_{\tilde{x} \in \mathcal{X}^\star(\tilde{\boldsymbol{\lambda}})} \|x - \tilde{x}\|_\infty. \tag{6}$$

Specifically, the following result shows that both the supremum and the infimum are attained, and that g_η is continuous over $\mathcal{X} \times \mathcal{M}$.

Lemma 19. *Let $\eta > 0$. Consider $g_\eta : \mathcal{X} \times \mathcal{M} \to \mathbb{R}$ defined as per Equation (6). Then, g_η is continuous over $\mathcal{X} \times \mathcal{M}$. Furthermore, both the* sup *and the* inf *are attained.*

Proof. First, consider $\inf_{\tilde{x} \in \mathcal{X}^\star(\tilde{\boldsymbol{\lambda}})} \|x - \tilde{x}\|_\infty$. This infimum is attained for all $\boldsymbol{\lambda} \in \Theta$ since it is an infimum over a compact set of a continuous function. Furthermore, by Theorem 5, the function is continuous. Therefore, the supremum is attained as well.

Finally, consider $g_\eta(x, \boldsymbol{\lambda})$. Using the fact that the infimum is continuous and that $\{\tilde{\boldsymbol{\lambda}} \in \mathcal{M} : \|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \eta\}$ is a continuous correspondence of $\boldsymbol{\lambda}$ (Lemma 4), we can apply Theorem 5, and we have that g_η is continuous over its domain, thus concluding the proof. □

We continue by optimizing $g_\eta(x, \boldsymbol{\lambda})$ in its second argument, *i.e.*,:

$$r_\eta(x) = \inf_{\boldsymbol{\lambda} \in \mathrm{cl}(\neg x)} g_\eta(x, \boldsymbol{\lambda}).$$

As above, we show that the infimum is attained, and that r_η is continuous over $\mathcal{X}$. This proof requires an extended version of the Berge's maximum theorem that can handle correspondences which are only lower hemicontinuous (see Theorem 6). Indeed, as we proved in Lemma 16, we only have a lower hemicontinuity result on the correspondence $x \rightrightarrows \text{cl}(\neg x)$.

Lemma 20. *Let $\eta > 0$. Let $r_\eta : \mathcal{X} \to \mathbb{R}$ be defined as follows:*

$$r_\eta(x) = \inf_{\boldsymbol{\lambda} \in \text{cl}(\neg x)} g_\eta(x, \boldsymbol{\lambda}).$$

Then, $r_\eta(x)$ is continuous over $\mathcal{X}$. Furthermore, the infimum is attained.

Proof. Let $\mathbb{X} = \mathcal{X}$, $\mathbb{Y} = \mathcal{M}$, $\phi(x) = \text{cl}(\neg x)$ for all $x \in \mathbb{X}$, and

$$u(x, \boldsymbol{\lambda}) = \max_{\tilde{\boldsymbol{\lambda}} \in \mathcal{M} : \|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \eta} \min_{\tilde{x} \in \mathcal{X}^\star(\tilde{\boldsymbol{\lambda}})} \|x - \tilde{x}\|.$$

We want to apply Theorem 6, which extends Berge's maximum theorem for lower hemicontinous correspondences, to prove the continuity of r_η. Before proceeding, we invite the reader to consult Appendix H.4 for useful definitions that will be used within this proof.

Now, we first note that, due to Lemma 16, $\phi(x)$ is lower hemicontinuous over $\mathbb{X}$.

Thus, to apply Theorem 6, it remains to check that u is $\mathbb{K}$-inf-compact and upper semicontinuous on $\text{Gr}_{\mathbb{X}}(\phi)$. First, we note that u is continuous on $\mathbb{X} \times \mathbb{Y}$ (Lemma 19), and hence upper semicontinuous. Furthermore, let $K \subset \mathbb{X}$ be compact and consider $\text{Gr}_K(\phi)$. Then, let us analyze the level sets $\mathcal{D}_u(\alpha, \text{Gr}_K(\phi))$ for $\alpha \in \mathbb{R}$. Then, $\mathcal{D}_u(\alpha, \text{Gr}_K(\phi)) \subseteq \mathbb{X} \times \mathbb{Y}$, which is compact, hence bounded. Thus, $\mathcal{D}_u(\alpha, \text{Gr}_K(\phi))$ is bounded as well. Moreover, $\mathcal{D}_u(\alpha, \text{Gr}_K(\phi))$ is closed since it is the preimage of a closed set of a continuous function. We have thus proved that u is $\mathbb{K}$-inf-compact and upper semicontinuous on $\text{Gr}_{\mathbb{X}}(\phi)$. From Theorem 6, r_η is continuous over $\mathbb{X}$

Finally, the supremum can be replaced by the maximum since the objective function is continuous and the optimization domain is compact. □

Finally, we optimize $r_\eta(x)$ over the possible answers $x \in \mathcal{X}$. The resulting value, *i.e.*, κ_η, will be than used to define $\bar{\rho}_\eta$. Specifically, in Lemma 22, we will set $\kappa_\eta = \bar{\rho}_\eta/2$. That being said, in the following lemma we show that the infimum of $r_\eta(x)$ is attained at some point $x \in \mathcal{X}$. Furthermore, it also shows that $\kappa_\eta > 0$.

Lemma 21. *Let $\eta > 0$, and let $\kappa_\eta = \inf_{x \in \mathcal{X}} r_\eta(x)$. The infimum is attained,* i.e., *$\kappa_\eta = \min_{x \in \mathcal{X}} r_\eta(x)$. Furthermore, $\kappa_\eta > 0$.*

Proof. First, we note that the infimum is attained. Indeed, from Lemma 20, r_η is continuous. The optimization domain is compact, and, therefore, the infimum is attained.

Secondly, we want to prove that $\kappa_\eta > 0$. Given the definition of κ_η, and since the infimum is attained, this is equivalent to proving that, for all $x \in \mathcal{X}$, $r_\eta(x) > 0$. From Lemma 20, we know that the min over the $\text{cl}(\neg x)$ is attained, therefore, we want to prove that for all $x \in \mathcal{X}, \boldsymbol{\lambda} \in \text{cl}(\neg x)$, $g_\eta(x, \boldsymbol{\lambda}) > 0$.

In other words, consider $\text{Gr}_{\mathcal{X}}(\text{cl}\neg)$ and let $(x, \boldsymbol{\lambda}) \in \text{Gr}_{\mathcal{X}}(\text{cl}\neg)$. We need to prove that $g_\eta(x, \boldsymbol{\lambda}) > 0$. We proceed by cases.

- If $(x, \boldsymbol{\lambda})$ is an isolated point in $\text{Gr}_{\mathcal{X}}(\text{cl}\neg)$, then $\boldsymbol{\lambda}$ is an isolated point in $\text{cl}(\neg x)$ (by virtue of Lemma 17), and, hence, $\boldsymbol{\lambda} \in \neg x$ (as guaranteed by Lemma 18). Thus, we have that:

$$g_\eta(x, \boldsymbol{\lambda}) \geq \min_{\tilde{x} \in \mathcal{X}^\star(\boldsymbol{\lambda})} \|x - \tilde{x}\|_\infty > 0.$$

- If, on the other hand, $(x, \boldsymbol{\lambda})$ is a limit point, there are two sub-cases. Either $\boldsymbol{\lambda}$ is an isolated point of $\text{cl}(\neg x)$ or it is a limit point of $\text{cl}(\neg x)$. If $\boldsymbol{\lambda}$ is an isolated point of $\text{cl}(\neg x)$ we have that $\boldsymbol{\lambda} \in \neg x$ (thanks to Lemma 18), and we can proceed as for isolated points of $\text{Gr}_{\mathcal{X}}(\text{cl}\neg)$.

- If $\boldsymbol{\lambda}$ is a limit point of both the graph and of $\mathrm{cl}(\neg x)$, there are two further cases. Either $\boldsymbol{\lambda} \in \neg x$, and we can proceed as for isolated points of $\mathrm{Gr}_{\mathcal{X}}(\mathrm{cl}\neg)$, or $\boldsymbol{\lambda} \in \mathrm{cl}(\neg x)$ and $\boldsymbol{\lambda} \notin \neg x$. In this last case, however, since $\boldsymbol{\lambda}$ is a limit point of $\mathrm{cl}(\neg x)$, there exists a sequence $\{\boldsymbol{\lambda}_n\}_{n\in\mathbb{N}}$ such that (i) $\boldsymbol{\lambda}_n \neq \boldsymbol{\lambda}$, $\forall n \in \mathbb{N}$, (ii) $\boldsymbol{\lambda}_n \in \neg x$, $\forall n \in \mathbb{N}$, and (iii) $\{\boldsymbol{\lambda}_n\} \to \boldsymbol{\lambda}$. Therefore, from (ii)+(iii), we have that for $\eta > 0$, there exists n_η such that, for all $n \geq n_\eta$, $\|\boldsymbol{\lambda}_n - \boldsymbol{\lambda}\|_\infty \leq \eta$ and $\boldsymbol{\lambda}_n \in \neg x$. Consider, *e.g.*, $\boldsymbol{\lambda}_{n_\eta}$; then, we have that:

$$g_\eta(x, \boldsymbol{\lambda}) = \sup_{\tilde{\boldsymbol{\lambda}} \in \mathcal{M}: \|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \eta} \inf_{\tilde{x} \in \mathcal{X}^\star(\tilde{\boldsymbol{\lambda}})} \|x - \tilde{x}\|_\infty \geq \inf_{\tilde{x} \in \mathcal{X}^\star(\boldsymbol{\lambda}_{n_\eta})} \|x - \tilde{x}\|_\infty > 0,$$

  thus leading to the desired result.

Therefore, all $(x, \boldsymbol{\lambda}) \in \mathrm{Gr}_{\mathcal{X}}(\mathrm{cl}\neg)$, we have that $g_\eta(x, \boldsymbol{\lambda}) > 0$ and therefore $\kappa_\eta > 0$. □

Now, we show that for all $\rho \in (0, \kappa_\eta]$, $\forall \bar{x} \in \mathcal{X}, \boldsymbol{\lambda} \in \mathrm{cl}(\neg \bar{x})$ there exists $\tilde{\boldsymbol{\lambda}} \in \neg \mathcal{B}_\rho(\bar{x})$ such that $\|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \eta$. As we shall see, combining this result with Lemma 23, leads to the desired result.

Lemma 22. *For all $\eta > 0$ sufficiently small, there exists $\bar{\rho}_\eta$ such that, for all $\rho \in (0, \bar{\rho}_\eta]$, it holds that:*

$$\forall \bar{x} \in \mathcal{X}, \boldsymbol{\lambda} \in \mathrm{cl}(\neg \bar{x}),\ \exists \tilde{\boldsymbol{\lambda}} \in \neg \mathcal{B}_\rho(\bar{x}) : \|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \eta.$$

Proof. Consider any $x \in \mathcal{X}$ and $\boldsymbol{\lambda} \in \mathrm{cl}(\neg x)$. Let $\bar{\boldsymbol{\lambda}}_{x,\boldsymbol{\lambda}}$ be defined as:

$$\bar{\boldsymbol{\lambda}}_{x,\boldsymbol{\lambda}} \in \operatorname*{argmax}_{\tilde{\boldsymbol{\lambda}} \in \mathcal{M}: \|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \eta} \min_{\tilde{x} \in \mathcal{X}^\star(\tilde{\boldsymbol{\lambda}})} \|x - \tilde{x}\|.$$

From Lemma 19, $\bar{\boldsymbol{\lambda}}_{x,\boldsymbol{\lambda}}$ is well-defined. Furthermore, by definition $\|\bar{\boldsymbol{\lambda}}_{x,\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \eta$.

Take $\bar{\rho}_\eta = \kappa_\eta/2$, and let $\rho \leq \bar{\rho}_\eta$. In the following, we prove that $\bar{\boldsymbol{\lambda}}_{x,\boldsymbol{\lambda}} \in \neg \mathcal{B}_\rho(x)$.

$$\begin{aligned}
\neg \mathcal{B}_\rho(x) &\supseteq \neg \mathcal{B}_{\kappa_\eta/2}(x) \\
&= \{\boldsymbol{\theta} \in \mathcal{M} : \forall \tilde{x} \in \mathcal{B}_{\kappa_\eta/2}(x), \tilde{x} \notin \mathcal{X}^\star(\boldsymbol{\theta})\} \\
&\supseteq \{\bar{\boldsymbol{\lambda}}_{x,\boldsymbol{\lambda}}\}
\end{aligned}$$

where the last step can be proved by contradiction. Suppose it is false, *i.e.*, there exists $\bar{x} \in \mathcal{B}_{\kappa_\eta/2}(x)$ such that $\bar{x} \in \mathcal{X}^\star(\bar{\boldsymbol{\lambda}}_{x,\boldsymbol{\lambda}})$. Then, it holds that:

$$\|x - \bar{x}\|_\infty \leq \frac{\kappa_\eta}{2} \leq \frac{1}{2} \min_{\tilde{x} \in \mathcal{X}^\star(\bar{\boldsymbol{\lambda}}_{x,\boldsymbol{\lambda}})} \|x - \tilde{x}\|_\infty \leq \frac{1}{2}\|x - \bar{x}\|_\infty < \|x - \bar{x}\|_\infty,$$

where in the last step we used $\kappa_\eta > 0$. This leads to a contradiction and concludes the proof. □

Now, we show that, when Lemma 22 holds, then Assumption 3 holds as well.

Lemma 23 (Sufficient condition for Assumption 3)**.** *If, for all $\tilde{\epsilon} > 0$, there exists $\rho > 0$, such that:*

$$\forall \bar{x} \in \mathcal{X}, \boldsymbol{\lambda} \in \mathrm{cl}(\neg \bar{x}),\ \exists \tilde{\boldsymbol{\lambda}} \in \neg \mathcal{B}_\rho(\bar{x}) : \|\tilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}\|_\infty \leq \tilde{\epsilon}.$$

then Assumption 3 holds.

Proof. Let $(\boldsymbol{\mu}, \boldsymbol{\omega}) \in \Theta^K \times \Delta_K$. Let $\boldsymbol{\lambda}^* \in \mathrm{argmin}_{\boldsymbol{\lambda} \in \mathrm{cl}(\neg x)} \sum_{k \in [K]} \omega_k d(\mu_k, \lambda_k)$ and denote by $\boldsymbol{\theta}^{\boldsymbol{\lambda}^*} \in \neg \mathcal{B}_\rho(x)$ be such that $\|\boldsymbol{\lambda}^* - \boldsymbol{\theta}^{\boldsymbol{\lambda}^*}\|_\infty \leq \tilde{\epsilon}$.[17] Let $\tilde{\boldsymbol{\lambda}}^*$ and $\tilde{\boldsymbol{\theta}}^{\boldsymbol{\lambda}^*}$ be the natural parameters of the distributions related to means

[17] Note that $\boldsymbol{\lambda}^*$ is well defined. Indeed, $\mathrm{cl}(\neg x)$ is compact, since $\mathcal{M}$ is compact. Furthermore, the function that is being optimized is continuous in $\boldsymbol{\lambda}$, and thus the min is attained.

$\boldsymbol{\lambda}^*$ and $\boldsymbol{\theta}^{\boldsymbol{\lambda}^*}$, respectively. Then, it holds that:

$$
\begin{aligned}
D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg\mathcal{B}_\rho(x)) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x) &= \inf_{\boldsymbol{\lambda} \in \neg\mathcal{B}_\rho(x)} \sum_{k \in [K]} \omega_k d(\mu_k, \lambda_k) - \inf_{\boldsymbol{\lambda} \in \neg x} \sum_{k \in [K]} \omega_k d(\mu_k, \lambda_k) \\
&\leq \inf_{\boldsymbol{\lambda} \in \neg\mathcal{B}_\rho(x)} \sum_{k \in [K]} \omega_k d(\mu_k, \lambda_k) - \inf_{\boldsymbol{\lambda} \in \mathrm{cl}(\neg x)} \sum_{k \in [K]} \omega_k d(\mu_k, \lambda_k) && \text{(Since } \neg x \subseteq \mathrm{cl}(\neg x)\text{)} \\
&\leq \sum_{k \in [K]} \omega_k \left(d(\mu_k, \theta_k^{\boldsymbol{\lambda}^*}) - d(\mu_k, \lambda_k^*) \right) && \text{(Since } \boldsymbol{\theta}^{\boldsymbol{\lambda}^*} \in \neg\mathcal{B}_\rho(x) \text{ and by definition of } \boldsymbol{\lambda}^*\text{)} \\
&\leq \max_{k \in [K]} \left| d(\mu_k, \theta_k^{\boldsymbol{\lambda}^*}) - d(\mu_k, \lambda_k^*) \right| \\
&= \max_{k \in [K]} \left| d(\lambda_k^*, \theta_k^{\boldsymbol{\lambda}^*}) + (\tilde{\theta}_k^{\boldsymbol{\lambda}^*} - \tilde{\lambda}_k^*)(\mu_k - \lambda_k^*) \right| && \text{(Lemma 44)} \\
&\leq \max_{k \in [K]} \left| C_2 (\lambda_k^* - \theta_k^{\boldsymbol{\lambda}^*})^2 + M(\tilde{\theta}_k^{\boldsymbol{\lambda}^*} - \tilde{\lambda}_k^*) \right| && \text{(Corollary 2 and } \Theta \text{ bounded)} \\
&\leq \max_{k \in [K]} \left(C_2 M \left| \lambda_k^* - \theta_k^{\boldsymbol{\lambda}^*} \right| + C_1 M \left| \lambda_k^* - \theta_k^{\boldsymbol{\lambda}^*} \right| \right) && \text{(Corollary 2 and } \Theta \text{ bounded)} \\
&= (C_1 + C_2) M ||\boldsymbol{\lambda}^* - \boldsymbol{\theta}^{\boldsymbol{\lambda}^*}||_\infty .
\end{aligned}
$$

Here, since Θ is bounded and contained in an open interval, we have used $(\mu - \mu') \leq M$ for any $\mu, \mu' \in \Theta$, where $M \coloneqq \max_{\mu \in \Theta} \mu - \min_{\mu \in \Theta} \mu$. Now, taking $\tilde{\epsilon} < \frac{\epsilon}{M(C_1 + C_2)}$, concludes the proof. □

Theorem 1 (Continuous Correspondence Implies Assumption 3)**.** *If $\boldsymbol{\mu} \rightrightarrows \mathcal{X}^\star(\boldsymbol{\mu})$ is continuous, and $\mathcal{M}$ and $\mathcal{X}$ are compact sets, then, Assumption 3 holds.*

Proof. Combine Lemma 23 together with Lemma 22. □

C Proof of the Lower Bound

In this section, we first provide a sketch of the proof that outlines all the underlying ideas and the differences with respect to previous works, and then we present the formal arguments.

C.1 Proof Sketch of Theorem 2

The general idea behind the proof is inspired by the lower bound for multiple answers problems presented in Degenne and Koolen [2019]. Specifically, in Degenne and Koolen [2019], the authors start by noticing that for any $T \in \mathbb{N}$, using Markov's inequality, one has that:

$$
\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta] = T(1 - \mathbb{P}_{\boldsymbol{\mu}}(\tau_\delta \leq T) \geq T \left(1 - \left(\delta + \sum_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} \mathbb{P}_{\boldsymbol{\mu}} \left(\{\tau_\delta \leq T\} \text{ and } \{\hat{x}_{\tau_\delta} = x\} \right) \right) \right) .
$$

Then, the proof follows by upper bounding $\mathbb{P}_{\boldsymbol{\mu}} \left(\{\tau_\delta \leq T\} \text{ and } \{\hat{x}_{\tau_\delta} = x\} \right)$ for each $x \in \mathcal{X}^\star(\boldsymbol{\mu})$ using change-of-measure arguments. As one can see, however, such an argument can actually be applied only when $\mathcal{X}^\star(\boldsymbol{\mu})$ is finite and, thus, complications arise in our infinite answer setting.

To solve this issue and prove Theorem 2, we combine three distinct elements, that are:

(i) An exact covering of the set $\mathcal{X}^\star(\boldsymbol{\mu})$ using balls of radius ρ

(ii) A change-of measure arguments that are directly related to this cover and to our new extended notion of alternative models over sets

(iii) A limit argument for $\rho \to 0$

As we now discuss, these three ingredients allows us to "reduce" the infinite answer problem to a finite answer one.

First, we note that, since $\mathcal{X}^\star(\boldsymbol{\mu})$ is compact, it admits a finite cover $\{\widetilde{\mathcal{X}}_i\}_{i=1}^{n_\rho}$ of $n_\rho \in \mathbb{N}$ elements using sets $\widetilde{\mathcal{X}}_i$ which are inscribed in balls of radius ρ (Lemma 43). Now, the idea is to fix a cover and try to follow the arguments of Degenne and Koolen [2019]. Specifically, we are going apply change of measure arguments by directly exploiting the sets $\{\widetilde{\mathcal{X}}_i\}_i$. Thus, for any $T \in \mathbb{N}$, using Markov's inequality, one has that:

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta] \geq T\left(1 - \mathbb{P}_{\boldsymbol{\mu}}(\tau_\delta \leq T)\right) \geq T\left(1 - \left(\delta + \sum_{i=1}^{n_\rho} \mathbb{P}_{\boldsymbol{\mu}}\left(\{\tau_\delta \leq T\} \text{ and } \{\hat{x}_{\tau_\delta} \in \widetilde{\mathcal{X}}_i\}\right)\right)\right),$$

where in the second step we have used the δ-correctness of the algorithm on regions that are complementary to $\mathcal{X}^\star(\boldsymbol{\mu})$.

Let $\mathcal{E}_i = \{\{\tau_\delta \leq T\} \text{ and } \{\hat{x}_{\tau_\delta} \in \widetilde{\mathcal{X}}_i\}\}$. From here, the idea is to upper bound $\mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}_i)$. Using change-of-measure arguments (Lemma 25), we can relate this probability to the one of the same event but under models in $\neg\widetilde{\mathcal{X}}_i$, *i.e.*, models for which *all* answers within $\widetilde{\mathcal{X}}_i$ are not correct. In this sense, the definition of alternative sets over sets of answers plays a crucial role, as it allows to obtain that, for all $\beta > 0$, some problem dependent constant α, and some $\boldsymbol{\lambda} \in \neg\widetilde{\mathcal{X}}_i$:

$$\begin{aligned}\mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}_i) &\leq \exp\left(TD(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_i) + \beta\right)\mathbb{P}_{\boldsymbol{\lambda}}(\mathcal{E}_i) + \exp\left(\frac{-\beta^2}{2T\alpha}\right)\\ &\leq \exp\left(TD(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_i) + \beta\right)\delta + \exp\left(\frac{-\beta^2}{2T\alpha}\right),\end{aligned}$$

where in the second step, we explicitly use $\boldsymbol{\lambda} \in \neg\widetilde{\mathcal{X}}_i$. For that step, indeed, it is required that every answer within $\widetilde{\mathcal{X}}_i$ is not a correct one for $\boldsymbol{\lambda}$. From here, one can use standard arguments from Degenne and Koolen [2019], and obtain (Lemma 26) the following asymptotic result:

$$\liminf_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} \geq \frac{1}{\max_{i \in [n_\rho]} D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_i)}, \tag{7}$$

The proof of Theorem 2 then follows by analyzing Equation (7) as $\rho \to 0$.

C.2 Proof of Theorem 2

We start the proof by introducing some preliminary results.

Lemma 24 (Minimax Results)**.** *Let $\boldsymbol{\mu} \in \Theta^K$ and $\Lambda \subset \mathcal{M}$ then*

$$D(\boldsymbol{\mu}, \Lambda) = \inf_{\mathbb{P}} \max_{k \in [K]} \mathbb{E}_{\boldsymbol{\lambda} \sim \mathbb{P}}[d(\mu_k, \lambda_k)],$$

where the infimum ranges over probability distributions on Λ supported on (at most) K points.

Proof. This result is a direct consequence of Lemma 2 in Degenne et al. [2019]. In Degenne et al. [2019] the authors state the result for a set $D(\boldsymbol{\mu}, \neg x)$, but it actually holds for any set Λ. □

We say that a distribution $q \in \Delta_K$, supported on $\boldsymbol{\lambda}^1, \ldots, \boldsymbol{\lambda}^K$, is optimal for $D(\boldsymbol{\mu}, \Lambda)$ for a given $\Lambda \subseteq \mathcal{M}$, if it attains the infimum of Lemma 24, *i.e.*, if

$$D(\boldsymbol{\mu}, \Lambda) = \max_{k \in [K]} \sum_{j \in [K]} q_j d(\mu_k, \lambda_k^j).$$

In particular, we are interested in the case in which $\Lambda = \neg\widetilde{\mathcal{X}}$, for some $\widetilde{\mathcal{X}}$ such that $\neg\widetilde{\mathcal{X}} \neq \emptyset$. In the following, we will assume, as in Degenne and Koolen [2019], that the infimum in $D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}})$ is attained in $\neg\widetilde{\mathcal{X}}$. If not, one can apply the following arguments to a sequence of ϵ-optimal distributions and let $\epsilon \to 0$.

Given this consideration, we recall a relevant change of measure arguments that have been previously used in BAI, *i.e.*, Lemma 19 of Degenne and Koolen [2019]. Before doing that, we introduce some necessary notation. For any model $\boldsymbol{\mu} \in \mathcal{M}$ and any k, we denote with $\tilde{\mu}_k$ the natural parameter of a distribution of the exponential family with mean μ_k. Further details on canonical exponential families are deferred to Appendix H.2.

Lemma 25 (Change of Measure)**.** *Fix $\boldsymbol{\mu} \in \mathcal{M}$ and let $\widetilde{\mathcal{X}} \subseteq \mathcal{X}$ be a subset of answers such that $\neg\widetilde{\mathcal{X}} \neq \emptyset$. Let $\boldsymbol{\lambda}^1, \dots, \boldsymbol{\lambda}^K$ and $\boldsymbol{q} \in \Delta_K$ be an optimal distribution for $D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}})$. Let $\alpha_k := \tilde{\mu}_k - \sum_{j\in[K]} q_j \tilde{\lambda}^j_k$ and $\bar{\alpha} = \max_{k\in[K]} \alpha_k$. Fix a sample size t and any event $\mathcal{E} \in \mathcal{F}_t$. Then, for any $\beta > 0$ it holds that:*

$$\max_{k\in[K]} \mathbb{P}_{\boldsymbol{\lambda}^k}(\mathcal{E}) \geq \exp\left\{-tD(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}) - \beta\right\}\left(\mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}) - \exp\left\{\frac{-\beta^2}{2t\bar{\alpha}^2}\right\}\right).$$

Proof. The proof is as in Degenne and Koolen [2019, Lemma 19]. □

Lemma 25 can be interpreted as follow. Whenever $t \ll D(\mu, \neg\widetilde{\mathcal{X}})^{-1}$, then if $\mathcal{E}$ is likely under $\boldsymbol{\mu}$ than it must also be likely under at least one $\boldsymbol{\lambda}^k$.

At this point, we derive the following intermediate results. The proof scheme is inspired by Theorem 1 in Degenne and Koolen [2019]. The key difference is that now we are applying those arguments to infinite answer identification problems, where each $\mathcal{X}^\star(\boldsymbol{\mu})$ is an arbitrary compact set.

Lemma 26 (Intermediate Result)**.** *For every $\boldsymbol{\mu} \in \mathcal{M}$ and for any $\rho > 0$ sufficiently small, there exists a finite set of answers $\{x_j\}_{j=1}^{n_\rho}$ and $x_j \in \mathcal{X}^\star(\boldsymbol{\mu})$ such that, if we define $\widetilde{\mathcal{X}}_j = \mathcal{B}_\rho(x_j) \cap \mathcal{X}^\star(\boldsymbol{\mu})$, then $\mathcal{X}^\star(\boldsymbol{\mu}) = \bigcup_{j=1}^{n_\rho} \widetilde{\mathcal{X}}_j$. Moreover, there exists $\widetilde{\mathcal{X}}_j$ in the cover such that $D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j) > 0$. Furthermore, for any δ-correct algorithm it holds that:*

$$\liminf_{\delta\to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} \geq \min_{j\in[n_\rho]: D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)>0} D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)^{-1}.$$

Proof. Fix $T > 0$ (to be defined later), due to the Markov's inequality, we have that:

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta] \geq T(1 - \mathbb{P}_{\boldsymbol{\mu}}(\tau_\delta \leq T)). \tag{8}$$

Before analyzing $\mathbb{P}_{\boldsymbol{\mu}}(\tau_\delta \leq T)$, we recall that $\mathcal{X}^\star(\boldsymbol{\mu})$ is compact. Therefore, by Lemma 43, we know that, for any $\rho > 0$, there exists a finite collection of compact subsets $\{\widetilde{\mathcal{X}}_j\}_{j\in[N_\rho]}$ such that $\mathcal{X}^\star(\boldsymbol{\mu}) = \bigcup_{j=1}^{N_\rho} \widetilde{\mathcal{X}}_j = \bigcup_{j=1}^{N_\rho} (\mathcal{X}^\star(\boldsymbol{\mu}) \cap \mathcal{B}_\rho(x_j))$ for some $x_j \in \mathcal{X}^\star(\boldsymbol{\mu})$. Furthermore, from Assumption 3, we also know that there exists $\bar{\rho}$ sufficiently small such that, for all $\rho \leq \bar{\rho}$, $\neg\mathcal{B}_\rho(x) \neq \emptyset$ for all $x \in \mathcal{X}$. In the following, we thus consider any ρ that satisfies this property.

Now, from Assumption 2 we know that there exist $\bar{x}$ such that $\boldsymbol{\mu} \notin \mathrm{cl}(\neg\bar{x})$. We add the set $\mathcal{X}^\star(\boldsymbol{\mu}) \cap \mathcal{B}_\rho(\bar{x})$ to the aforementioned cover, thus obtaining a cover of size $n_\rho = N_\rho + 1$, such that each element of the cover is of the kind $\mathcal{X}^\star(\boldsymbol{\mu}) \cap \mathcal{B}_\rho(x_j)$ for some $x_j \in \mathcal{X}^\star(\boldsymbol{\mu})$ and $\neg\mathcal{B}_\rho(x_j) \neq \emptyset$ for all $j \in [n_\rho]$. Furthermore, by applying Lemma 14, we know there exists x_j (*i.e.*, $\bar{x}$) for some $j \in [n_\rho]$ such that $D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j) > 0$.[18] As we shall see in a few steps, this property will be used to invert $\max_{j\in[n_\rho]} D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)$.

Now, we analyze $\mathbb{P}_{\boldsymbol{\mu}}(\tau_\delta \leq T)$. Let us introduce the event $\mathcal{E} = \{\hat{x}_{\tau_\delta} \in \mathcal{X}^\star(\boldsymbol{\mu})\}$ and the events $\mathcal{E}_j = \{\tau_\delta \leq T\} \cap \{\hat{x}_{\tau_\delta} \in \widetilde{\mathcal{X}}_j\}$ for all $j \in [n_\rho]$.[19] Consider the following:

$$\begin{aligned}
\mathbb{P}_{\boldsymbol{\mu}}(\tau_\delta \leq T) &= \mathbb{P}_{\boldsymbol{\mu}}(\tau_\delta \leq T|\mathcal{E})\mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}) + \mathbb{P}_{\boldsymbol{\mu}}(\tau_\delta \leq T|\mathcal{E}^c)\mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}^c) \\
&\leq \mathbb{P}_{\boldsymbol{\mu}}(\{\tau_\delta \leq T\} \cap \mathcal{E}) + \delta \\
&\leq \sum_{j=1}^{n_\rho} \mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}_j) + \delta,
\end{aligned} \tag{9}$$

[18] Indeed, since $\boldsymbol{\mu} \notin \mathrm{cl}(\neg\bar{x})$, it also holds that $\boldsymbol{\mu} \notin \widetilde{\mathcal{X}}_j$.

[19] We note that $\mathcal{E}_j$ is measurable since $\hat{x}_{\tau_\delta}$ is measurable with respect to $\mathcal{F}_{\tau_\delta}$ and $\widetilde{\mathcal{X}}_j$ is a Borel set.

where in the first step, we used the law of total probability; in the second one, we used the fact that the algorithm is δ-correct; and in the third one we applied a union bound.

At this point, consider all $j \in [n_\rho]$, and analyze $\mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}_j)$. Let $\boldsymbol{\lambda}^1, \dots, \boldsymbol{\lambda}^K$ and $\boldsymbol{q}$ be an optimal distribution according to $D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)$. Then, from Lemma 25, we have, for all $\beta > 0$:

$$\begin{aligned}\mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}_j) &\leq \exp\left(TD(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j) + \beta\right) \max_{k\in[K]} \mathbb{P}_{\boldsymbol{\lambda}^k}(\mathcal{E}_j) + \exp\left(\frac{-\beta^2}{2T\bar{\alpha}^2}\right) \\ &\leq \delta \exp\left(TD(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j) + \beta\right) + \exp\left(\frac{-\beta^2}{2T\bar{\alpha}^2}\right),\end{aligned}$$

where, in the second step, we have used that $\mathbb{P}_{\boldsymbol{\lambda}^k}(\mathcal{E}_j) \leq \delta$, since for all $\boldsymbol{\lambda}^k$ we have that $\boldsymbol{\lambda}^k \in \neg\widetilde{\mathcal{X}_j}$ and thus $\mathbb{P}_{\boldsymbol{\lambda}^k}(\mathcal{E}_j) \leq \delta$ due to the δ-correctness of the algorithm. Note that, for this step, the construction of the sets $\widetilde{\mathcal{X}}_j$ and the notion of extended alternative model plays a crucial role. Indeed, since $\widetilde{\mathcal{X}}_j = \{\boldsymbol{\lambda} \in \mathcal{M} : \forall x \in \widetilde{\mathcal{X}}_j, x \notin \mathcal{X}^\star(\boldsymbol{\mu})\}$ and since algorithms are δ-correct, then $\mathbb{P}_{\boldsymbol{\lambda}}(\mathcal{E}_j) \leq \mathbb{P}_{\boldsymbol{\lambda}}(\hat{x}_{\tau_\delta} \in \widetilde{\mathcal{X}}_j) \leq \delta$ for all $\boldsymbol{\lambda} \in \neg\widetilde{\mathcal{X}}_j$.

We now follow the argument of Degenne and Koolen [2019]. For a fixed $\eta \in (0,1)$, we define[20]

$$T = (1-\eta)\log(1/\delta) \min_{k\in[n_\rho]: D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)>0} D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)^{-1},$$

and take $\beta = \frac{\eta}{2\sqrt{1-\eta}}\sqrt{TD(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)\log(1/\delta)}$. Then

$$\mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}_j) \leq \delta^{\frac{\eta}{2}} + \delta^{\frac{\eta^2 D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}})}{8(1-\eta)\bar{\alpha}}},$$

which goes to zero for all $\eta > 0$ when $\delta \to 0$.

If we plug this into Equation (8) and using Equation (9) we obtain that

$$\begin{aligned}\frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} &\geq (1-\eta) \min_{k\in[n_\rho]: D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)>0} D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)^{-1} \left(1 - \delta - \sum_{j=1}^{n_\rho} \mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}_j)\right) \\ &\geq (1-\eta) \min_{k\in[n_\rho]: D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)>0} D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)^{-1} \left(1 - \delta - n_\rho\left(\delta^{\frac{\eta}{2}} + \delta^{\frac{\eta^2 D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}})}{8(1-\eta)\bar{\alpha}}}\right)\right)\end{aligned}$$

and thus, by taking the limit $\delta \to 0$, we obtain:

$$\lim_{\delta\to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} \geq (1-\eta) \min_{k\in[n_\rho]: D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)>0} D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)^{-1},$$

which holds for all $\eta > 0$ and thus proves the result. □

We are now ready to prove Theorem 2.

Theorem 2 (Lower Bound)**.** *For any $\boldsymbol{\mu} \in \mathcal{M}$, and any δ-correct algorithm it holds that:*

$$\liminf_{\delta\to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} \geq T^*(\boldsymbol{\mu}) = \frac{1}{D(\boldsymbol{\mu})}. \tag{1}$$

Proof of Theorem 2. Let $\rho > 0$ be sufficiently small. Then, by Lemma 26, we know that there exists a finite set of points $\{x_j\}_{j\in[n_\rho]} \subset \mathcal{X}^\star(\boldsymbol{\mu})$ such that there exists an exact cover of $\mathcal{X}^\star(\boldsymbol{\mu})$ of the form $\widetilde{\mathcal{X}}_j = \mathcal{X}^\star(\boldsymbol{\mu}) \cap \mathcal{B}_\rho(x_j)$ and such that for at least one j we have $D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j) > 0$ and

$$\liminf_{\delta\to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} \geq \min_{j\in[n_\rho]} D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)^{-1}.$$

[20] Note that the minimum here is finite since for at least one element of the cover has positive value.

First of all, we notice that $T^*(\boldsymbol{\mu}) \geq \min_{j\in[n_\rho]} D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)^{-1}$ always holds. Indeed, consider $x \in \operatorname{argmax}_{\tilde{x}\in\mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \neg\tilde{x})$ and consider any j such that $x \in \widetilde{\mathcal{X}}_j$. Then, we have that:

$$\begin{aligned} D(\boldsymbol{\mu}) &= \sup_{\boldsymbol{\omega}\in\Delta_K} \inf_{\boldsymbol{\lambda}\in\neg x} \sum_{k\in[K]} \omega_k d(\mu_k, \lambda_k) \\ &\leq \sup_{\boldsymbol{\omega}\in\Delta_K} \inf_{\boldsymbol{\lambda}\in\neg\widetilde{\mathcal{X}}_j} \sum_{k\in[K]} \omega_k d(\mu_k, \lambda_k) \\ &= D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j), \end{aligned}$$

where in the inequality step we have used $\neg\widetilde{\mathcal{X}}_j \subseteq \neg x$ since $x \in \widetilde{\mathcal{X}}_j$.

Then, let $\bar{j} \in \operatorname{argmax}_{j\in[n_\rho]} D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)$, and let $\bar{x}$ be such that $\widetilde{\mathcal{X}}_{\bar{j}} = \mathcal{X}^\star(\boldsymbol{\mu}) \cap \mathcal{B}_\rho(\bar{x})$ for some $\bar{x} \in \mathcal{X}^\star(\boldsymbol{\mu})$. Then, we have that:

$$\begin{aligned} \frac{1}{D(\boldsymbol{\mu})} - \min_{j\in[n_\rho]} \frac{1}{D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_j)} &= \frac{D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_{\bar{j}}) - D(\boldsymbol{\mu})}{D(\boldsymbol{\mu}) D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_{\bar{j}})} \\ &\leq \frac{D(\boldsymbol{\mu}, \neg\mathcal{B}_\rho(\bar{x})) - D(\boldsymbol{\mu})}{D(\boldsymbol{\mu}) D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_{\bar{j}})} && \text{(since } \neg\mathcal{B}_\rho(\bar{x}) \subseteq \neg\widetilde{\mathcal{X}}_{\bar{j}}) \\ &\leq \frac{D(\boldsymbol{\mu}, \neg\mathcal{B}_\rho(\bar{x})) - D(\boldsymbol{\mu}, \neg\bar{x})}{D(\boldsymbol{\mu}) D(\boldsymbol{\mu}, \neg\widetilde{\mathcal{X}}_{\bar{j}})} && \text{(since } D(\boldsymbol{\mu}) \geq D(\boldsymbol{\mu}, \neg x)) \\ &\leq \epsilon, \end{aligned}$$

where the last inequality comes from Assumption 3.[21] Letting $\epsilon \to 0$ concludes the proof. □

D Continuity Results

In this section, we provide results on the continuity of the different divergences involved in an infinite answer problem.

We begin with the following preliminary result.

Lemma 27. *For all $\boldsymbol{\mu} \in \Theta^K$, it holds that:*

$$D(\boldsymbol{\mu}) = \max_{x\in\mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \neg x) = \max_{x\in\mathcal{X}} D(\boldsymbol{\mu}, \neg x).$$

Proof. Let $x \in \mathcal{X}$ and $x \notin \mathcal{X}^\star(\boldsymbol{\mu})$. Then, consider any $\boldsymbol{\omega} \in \Delta_K$, we have that:

$$D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x) = \inf_{\boldsymbol{\lambda}\in\neg x} \sum_{k\in[K]} \omega_k d(\mu_k, \lambda_k) = 0. \tag{10}$$

Indeed, $\neg x = \{\boldsymbol{\lambda} \in \mathcal{M} : x \notin \mathcal{X}^\star(\boldsymbol{\lambda})\}$, and thus $\boldsymbol{\mu} \in \neg x$. Them, Equation (10) follows by $d(\mu, \mu) = 0$.

Observing that $D(\boldsymbol{\mu}, \neg x) \geq 0$ for all $x \in \mathcal{X}$ (i.e, Assumption 1) concludes the proof. □

Lemma 1 (Continuity)**.** *The following holds:*

(i) The function $(\boldsymbol{\mu}, \boldsymbol{\omega}, x) \to D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is continuous over $\Theta^K \times \Delta_K \times \mathcal{X}$.

(ii) The function $(\boldsymbol{\mu}, x) \to D(\boldsymbol{\mu}, \neg x)$ is continuous over $\Theta^K \times \mathcal{X}$ and $(\boldsymbol{\mu}, x) \rightrightarrows \boldsymbol{\omega}^*(\boldsymbol{\mu}, \neg x)$ is upper hemicontinuous and compact-valued.

(iii) The function $(\boldsymbol{\mu}, \boldsymbol{\omega}) \to \max_{x\in\mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is continuous over $\Theta^K \times \Delta_K$ and $(\boldsymbol{\mu}, \boldsymbol{\omega}) \rightrightarrows \operatorname{argmax}_{x\in\mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is upper hemicontinuous and compact-valued.

[21] Notice that, since Assumption 3 holds for any $\boldsymbol{\omega} \in \Delta_K$, it holds also for the supremum over $\boldsymbol{\omega}$. Indeed, $D(\boldsymbol{\mu}, \neg\mathcal{B}_\rho(\bar{x})) - D(\boldsymbol{\mu}, \neg\bar{x}) \leq D(\boldsymbol{\mu}, \bar{\boldsymbol{\omega}}, \neg\mathcal{B}_\rho(\bar{x})) - D(\boldsymbol{\mu}, \bar{\boldsymbol{\omega}}, \neg\bar{x})$ for any $\bar{\boldsymbol{\omega}}$ that attains the supremum in $D(\boldsymbol{\mu}, \neg\mathcal{B}_\rho(\bar{x}))$.

(iv) The function $\boldsymbol{\mu} \to D(\boldsymbol{\mu})$ is continuous over Θ^K. Moreover, $\boldsymbol{\mu} \rightrightarrows \boldsymbol{\omega}^\star(\boldsymbol{\mu})$ and $\boldsymbol{\mu} \rightrightarrows \mathcal{X}_F(\boldsymbol{\mu})$ are upper hemicontinuous and compact-valued over $\mathcal{S}$.

Proof. **i)** First, note that, for all $x \in \mathcal{X}$, the function $(\boldsymbol{\mu}, \boldsymbol{\omega}) \to D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is jointly continuous over $\Theta^K \times \Delta_K$. This is due to Degenne and Koolen [2019, Lemma 27]. Then, it remains to show that $(\boldsymbol{\mu}, \boldsymbol{\omega}, x) \to D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is jointly continuous over $\Theta^K \times \Delta_K \times \mathcal{X}$. Thus, for all $\boldsymbol{\mu}, \boldsymbol{\omega}, x \in \Theta^K \times \Delta_K \times \mathcal{X}$ and $\forall \epsilon > 0$ there exists $\kappa^\star = \kappa^\star_{\mu,\omega,x,\epsilon} > 0$ such that, for all $\boldsymbol{\mu}', \boldsymbol{\omega}', x' : \|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \kappa^\star, \|\boldsymbol{\omega} - \boldsymbol{\omega}'\|_\infty \leq \kappa^\star, \|x - x'\|_\infty \leq \kappa^\star$, it holds that $|D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x') - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)| \leq \epsilon$. Define $\kappa^\star = \min\{\tilde{\kappa}, \bar{\kappa}\}$, where $\bar{\kappa}$ and $\tilde{\kappa}$ are as follows. First, $\bar{\kappa}$ is such that:

$$D(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\omega}}, \neg \mathcal{B}_{\bar{\kappa}}(\bar{x})) - D(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\omega}}, \neg \bar{x}) \leq \frac{\epsilon}{2} \quad \forall \bar{x} \in \mathcal{X}, \bar{\boldsymbol{\mu}} \in \Theta^K, \bar{\boldsymbol{\omega}} \in \Delta_K. \tag{11}$$

Due to Assumption 3 such $\bar{\kappa}$ is guaranteed to exists. Secondly, $\tilde{\kappa} = \tilde{\kappa}_{\mu,\omega,x,\epsilon}$ is such that:

$$|D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)| \leq \frac{\epsilon}{2} \quad \forall \boldsymbol{\mu}' \in \mathcal{B}_{\tilde{\kappa}}(\boldsymbol{\mu}), \boldsymbol{\omega}' \in \mathcal{B}_{\tilde{\kappa}}(\boldsymbol{\omega}). \tag{12}$$

Such $\tilde{\kappa}$ is guaranteed to exists due to the continuity of $(\boldsymbol{\mu}, \boldsymbol{\omega}) \to D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ for any fixed $x \in \mathcal{X}$ due to Degenne and Koolen [2019, Lemma 27].

Then, we analyze $|D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x') - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)|$ by studying an upper bound on the sum of the following terms: $|D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x') - D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x)|$ and $|D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)|$. We start with the former. Suppose, $D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x') > D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x)$, then

$$\begin{aligned}
|D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x') - D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x)| &= D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x') - D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x) \\
&\leq D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg \mathcal{B}_{\bar{\kappa}}(x)) - D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x) \\
&\leq \frac{\epsilon}{2}.
\end{aligned}$$

where the first inequality holds since $x' \subseteq \mathcal{B}_{\kappa^\star}(x) \subseteq \mathcal{B}_{\bar{\kappa}}(x)$ and thus $\neg x' \supseteq \neg \mathcal{B}_{\bar{\kappa}}(x)$, and the second inequality follows from Equation (11). Equivalently, if $D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x') \leq D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x)$ then

$$\begin{aligned}
|D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x') - D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x)| &= D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x) - D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x') \\
&\leq D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg \mathcal{B}_{\bar{\kappa}}(x')) - D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x') \\
&\leq \frac{\epsilon}{2},
\end{aligned}$$

where the first inequality holds since $x \subseteq \mathcal{B}_{\kappa^\star}(x') \subseteq \mathcal{B}_{\bar{\kappa}}(x')$ and thus $\neg x \supseteq \neg \mathcal{B}_{\bar{\kappa}}(x')$, and the second inequality follows from Equation (11).

Similarly, for the second term, we have that:

$$|D(\boldsymbol{\mu}', \boldsymbol{\omega}', \neg x) - D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)| \leq \frac{\epsilon}{2},$$

by the fact that $\boldsymbol{\mu}' \in \mathcal{B}_{\kappa^\star}(\boldsymbol{\mu}), \boldsymbol{\omega}' \in \mathcal{B}_{\kappa^\star}(\boldsymbol{\omega})$ and Equation (12). This concludes the proof of the first statement.

The other three statements follow from various applications of Berge's maximum theorem (Theorem 5).

ii) $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is continuous over $\Theta^K \times \Delta_K \times \mathcal{X}$ and the maximization is over the simplex.

iii) The third claim is due to Berge's maximum theorem. Indeed, $\mathcal{X}^\star(\boldsymbol{\mu})$ is continuous and compact-valued and $D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is continuous. Hence, $(\boldsymbol{\mu}, \boldsymbol{\omega}) \to \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is continuous and $(\boldsymbol{\mu}, \boldsymbol{\omega}) \rightrightarrows \operatorname{argmax}_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is upper hemicontinuous and compact-valued.

iv) Finally, the last claim follows by applying the Berge's maximum theorem. We recall from Lemma 27 that $D(\boldsymbol{\mu}) = \max_{x \in \mathcal{X}} D(\boldsymbol{\mu}, \neg x)$. Since $D(\boldsymbol{\mu}, \neg x)$ is continuous and $\boldsymbol{\mu} \rightrightarrows \mathcal{X}$ is constant and compact-valued (Assumption 1), we obtain that $x_F(\boldsymbol{\mu})$ is upper hemicontinuous and compact-valued. Furthermore, since $\max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ is continuous and the simplex is a constant and compact set, we have that $\boldsymbol{\mu} \rightrightarrows \boldsymbol{\omega}^*(\boldsymbol{\mu})$ is upper hemicontinous and compact-valued. □

Corollary 1 (Uniform Continuity)**.** *Let $\mathcal{C} \subseteq \Theta^K$ and $\mathcal{H} \subseteq \mathcal{X}$ be compact sets. Then, we have that:*

- $(\boldsymbol{\omega}, \boldsymbol{\mu}, x) \to D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ *is uniformly continuous on* $\mathcal{C} \times \Delta_K \times \mathcal{H}$.
- $(\boldsymbol{\mu}, x) \to D(\boldsymbol{\mu}, \neg x)$ *is uniformly continuous on* $\mathcal{C} \times \mathcal{H}$.
- $(\boldsymbol{\mu}, \boldsymbol{\omega}) \to \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)$ *is uniformly continuous on* $\mathcal{C} \times \Delta_K$.
- $\boldsymbol{\mu} \to D(\boldsymbol{\mu})$ *is uniformly continuous on* $\mathcal{C}$

Proof. By Heine–Cantor, every continuous function is uniformly continuous on a compact domain. Then, apply Lemma 1. □

E Algorithm Analysis

We structure this section as follows. First, in Appendix E.1, we prove the correctness of the stopping and recommendation rules that we introduced in Section 5.1. Then, we present the analysis of Sticky Sequence TaS for any answer selection rule that ensures convergence (Appendix E.2). Finally, in Appendix E.3, we show how to build such converging sequences according to the properties of $\mathcal{X}$ and $\boldsymbol{\mu} \rightrightarrows \mathcal{X}^\star(\boldsymbol{\mu})$.

E.1 Correctness and Expected Stopping Time

We first recall the stopping and recommendation rules. Specifically, the stopping rule is as follows:

$$\tau_\delta = \inf \left\{ t \in \mathbb{N} : \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D\left(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x\right) > \beta_{t,\delta} \right\}. \tag{13}$$

Furthermore, it recommends:

$$\hat{x}_{\tau_\delta} \in \operatorname*{argmax}_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D\left(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x\right). \tag{14}$$

The following correctness analysis is based on concentration arguments that have been initially described by Ménard [2019], where the author is working under the assumption of Gaussian distributions with unitary variance. Specifically, the stopping threshold $\beta_{t,\delta}$ is set to:

$$\beta_{t,\delta} = \log\left(\frac{1}{\delta}\right) + K \log\left(4 \log\left(\frac{1}{\delta}\right) + 1\right) + 6K \log(\log(t) + 3) + K\tilde{C}, \tag{15}$$

where $\tilde{C}$ is a universal constant. All the results of this section can be extended (with a more complex notation) to distributions in any canonical exponential family by modifying appropriately $\beta_{t,\delta}$ (see e.g., Kaufmann and Koolen [2021]).

Lemma 28 (Correctness)**.** *For any sampling rule* $(A_t)_{t \geq 1}$*, the stopping and recommendation rules in Equations* (13)-(14) *guarantee that:*

$$\mathbb{P}_{\boldsymbol{\mu}}\left(\hat{x}_{\tau_\delta} \notin \mathcal{X}^\star(\boldsymbol{\mu})\right) \leq \delta.$$

Proof. It holds that:

$$\begin{aligned}
\mathbb{P}_{\boldsymbol{\mu}}\left(\hat{x}_{\tau_\delta} \notin \mathcal{X}^\star(\boldsymbol{\mu})\right) &= \mathbb{P}_{\boldsymbol{\mu}}\left(\exists t \in \mathbb{N} : \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D\left(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x\right) > \beta_{t,\delta} \text{ and } \hat{x}_t \notin \mathcal{X}^\star(\boldsymbol{\mu})\right) \\
&\leq \mathbb{P}_{\boldsymbol{\mu}}\left(\exists t \in \mathbb{N} \text{ and } x \notin \mathcal{X}^\star(\boldsymbol{\mu}) : D\left(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x\right) > \beta_{t,\delta}\right) \\
&\leq \mathbb{P}_{\boldsymbol{\mu}}\left(\exists t \in \mathbb{N} : \sum_{k \in [K]} N_k(t) d(\hat{\mu}_k(t), \mu_k) > \beta_{t,\delta}\right) \\
&\leq \delta,
\end{aligned}$$

where (i) in the first step we have used Equations (13)-(14), (ii) in the third one the fact that, since $x \notin \mathcal{X}^\star(\boldsymbol{\mu})$, then $\boldsymbol{\mu} \in \neg x$, and (iii) in the last one Proposition 1 in Ménard [2019]. □

The following results will be used to control the expected stoppping time of the proposed algorithms. Lemma 29 is a standard result; we include a proof for completeness.

Lemma 29. *Let $\{\mathcal{E}(t)\}_t$ be a sequence of $\mathcal{F}_t$-measurable events. Suppose there exists $T_0(\delta) \in \mathbb{N}$ such that, for all $t \geq T_0(\delta)$, $\mathcal{E}(t) \subseteq \{\tau_\delta < t\}$. Then, it holds that:*

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta] \leq T_0(\delta) + \sum_{t=0}^{+\infty} \mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}(t)^c).$$

Proof. By standard probabilistic arguments, we have that:

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta] = \sum_{t=0}^{+\infty} \mathbb{P}_{\boldsymbol{\mu}}(\tau_\delta > t) \leq T_0(\delta) + \sum_{t=T_0(\delta)}^{+\infty} \mathbb{P}_{\boldsymbol{\mu}}(\tau_\delta > t) \leq T_0(\delta) + \sum_{t=0}^{+\infty} \mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}(t)^c),$$

which concludes the proof. □

E.2 Sticky-Sequence Track-and-Stop

In this section, we prove Theorem 3, *i.e.*, we analyze Sticky-Sequence Track-and-Stop for any general answer selection rule that satisfies the fact that $\{x_t\}_t$ is converging to some $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$.

The analysis will be carried out under the good event $\mathcal{E}_t = \bigcap_{s=h(t)}^{t} \{\boldsymbol{\mu} \in C_s\}$ where $h(t) = \lceil \sqrt{t} \rceil$ and $C_s = \{\boldsymbol{\mu}' \in \mathcal{M} : D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \boldsymbol{\mu}') \leq \log(g(t))\}$ where $g(t) = Ct^{10}$ for some constant $C \in \mathbb{R}$. It is possible to show that, for an appropriate value of C, it holds that $\sum_{t=0}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}(t)^c)$ is finite (see Lemma 14 in Degenne and Koolen [2019]). Next, we first recall the definition of converging sequence $\{x_t\}_t$. Specifically, for all $\epsilon > 0$, there exists $T_\epsilon \in \mathbb{N}$ such that, for all $t \geq T_\epsilon$, on $\mathcal{E}(t)$, there exists $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu}) : \|x_s - \bar{x}\| \leq \epsilon$ for all $s \geq h(t)$.

Now that we clarified the setup, we first prove that, on $\mathcal{E}(t)$, C_t is shrinking toward $\boldsymbol{\mu}$. Note that this result is independent from Definition 1 and only depends on the definition of $\mathcal{C}_t$, $\mathcal{E}_t$ and the forced exploration of the algorithm (*i.e.*, the cumulative tracking procedure, Lemma 41).

Lemma 30. *For all $\epsilon > 0$, there exists T_ϵ such that, for all $t \geq T_\epsilon$, on $\mathcal{E}(t)$, $\|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \epsilon$ holds $\forall \boldsymbol{\mu}' \in C_s$ and all $s \geq h(t)$.*

Proof. For any $\boldsymbol{\mu}, \boldsymbol{\mu}'$, let $ch(\boldsymbol{\mu}, \boldsymbol{\mu}') = \inf_{\boldsymbol{\lambda} \in \mathbb{R}^K} \sum_{k \in [K]} (d(\lambda_k, \mu_k) + d(\lambda_k, \mu'_k))$.

Consider $T_\epsilon \geq \bar{T}$, where $\bar{T}$ is such that $\sqrt{h(\bar{T}) + K^2} - 2K > 0$ and let $t \geq T_\epsilon$. We recall that, by definition, on $\mathcal{E}(t)$, it holds that $\boldsymbol{\mu} \in C_s$ for all $s \geq h(t)$. Thus, $D(\hat{\boldsymbol{\mu}}(s), \boldsymbol{N}(s), \boldsymbol{\mu}) \leq \log(g(s))$ for all $s \geq h(t)$. Furthermore, for all $\boldsymbol{\mu}' \in C_s$, $D(\hat{\boldsymbol{\mu}}(s), \boldsymbol{N}(s), \boldsymbol{\mu}') \leq \log(g(s))$. Thus, on $\mathcal{E}(t)$, for all $s \geq h(t)$ and $\boldsymbol{\mu}' \in C_s$, we have that:

$$\sum_{k \in [K]} N_k(s) \left(d(\hat{\mu}_k(s), \mu_k) + d(\hat{\mu}_k(s), \mu'_k)\right) \leq 2 \log(g(s)),$$

Using Lemma 41 together with the fact that t is such that $\sqrt{h(t) + K^2} - 2K > 0$ (since $t \geq \bar{T}$), we obtain that, on $\mathcal{E}(t)$, $ch(\boldsymbol{\mu}, \boldsymbol{\mu}') \leq \frac{2 \log(g(s))}{\sqrt{s+K^2} - 2K}$, $\forall \boldsymbol{\mu}' \in \mathcal{C}_s$.

Now, observe that $\frac{2 \log(g(s))}{\sqrt{s+K^2} - 2K}$ is decreasing in s. It follows that, for all $\epsilon > 0$, there exists s_ϵ such that $\forall s \geq s_\epsilon$, $\|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \epsilon$. Indeed, let us analyze $ch(\boldsymbol{\mu}, \boldsymbol{\mu}')$. Using the sub-gaussianity of the arms together with the fact

that $ch(\boldsymbol{\mu}, \boldsymbol{\mu}') \leq \frac{2\log(g(s))}{\sqrt{s+K^2}-2K}$ we obtain:[22]

$$\begin{aligned}
\frac{2\log(g(s))}{\sqrt{s+K^2}-2K} &\geq ch(\boldsymbol{\mu}, \boldsymbol{\mu}') \\
&\geq \inf_{\boldsymbol{\lambda}\in\mathbb{R}^K} \sum_{k\in[K]} \frac{(\lambda_k-\mu_k)^2+(\lambda_k-\mu'_k)^2}{2\sigma^2} \\
&= \frac{1}{2\sigma^2}\sum_{k\in[K]} \left(\frac{\mu_k+\mu'_k}{2}-\mu_k\right)^2 + \left(\frac{\mu_k+\mu'_k}{2}-\mu'_k\right)^2 \\
&= \frac{1}{4\sigma^2}\sum_{k\in[K]} (\mu_k-\mu'_k)^2 \\
&= \frac{1}{4\sigma^2}\|\boldsymbol{\mu}-\boldsymbol{\mu}'\|_2^2 \\
&\geq \frac{1}{4\sigma^2}\|\boldsymbol{\mu}-\boldsymbol{\mu}'\|_\infty^2.
\end{aligned}$$

Thus, we have that, on $\mathcal{E}(t)$, for $s \geq h(t)$ and $\boldsymbol{\mu}' \in C_s$:

$$\|\boldsymbol{\mu}-\boldsymbol{\mu}'\|_\infty \leq \sqrt{\frac{8\sigma^2\log(g(s))}{\sqrt{s+K^2}-2K}}.$$

Let $T_\epsilon := \max\left\{\bar{T}, \inf\left\{n\in\mathbb{N} : \sqrt{\frac{8\log(g(h(n)))}{\sqrt{h(n)+K^2}-2K}} \leq \epsilon\right\}\right\}$. Then, for all $t \geq T_\epsilon$, $\|\boldsymbol{\mu}-\boldsymbol{\mu}'\|_\infty \leq \epsilon$ holds on $\mathcal{E}(t)$ for all $\boldsymbol{\mu}' \in C_s$ and $s \geq h(t)$. □

At this point, we combine Lemma 30 with Lemma 1 and Corollary 1 to prove the following result. Lemma 31 will be used to analyze the behavior of the algorithm in Lemma 32, which is at the core of our asymptotic optimality result.

Lemma 31. *If Definition 1 is satisfied, then, for all $\kappa > 0$ there exists $T_\kappa \in \mathbb{N}$ such that for all $t \geq T_\kappa$, on $\mathcal{E}(t)$, it holds that, for all $s \geq h(t)$ and $\boldsymbol{\mu}' \in C_s$:*

$$\begin{aligned}
&\left|D(\boldsymbol{\mu}) - D(\boldsymbol{\mu}')\right| \leq \kappa, \\
&\left|\max_{x\in\mathcal{X}_F(\boldsymbol{\mu})} D(\boldsymbol{\mu},\boldsymbol{\omega},\neg x) - \max_{x\in\mathcal{X}_F(\boldsymbol{\mu}')} D(\boldsymbol{\mu}',\boldsymbol{\omega},\neg x)\right| \leq \kappa \quad \forall \boldsymbol{\omega}\in\Delta_K.
\end{aligned}$$

Furthermore, there exists $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$ such that, on $\mathcal{E}(t)$:

$$\left|D(\boldsymbol{\mu},\boldsymbol{\omega},\neg\bar{x}) - D(\boldsymbol{\mu}',\boldsymbol{\omega},\neg x_s)\right| \leq \kappa, \quad \forall \boldsymbol{\omega}\in\Delta_K, \boldsymbol{\mu}'\in C_s, s\geq h(t).$$

Proof. First, from Corollary 1, we have that for all $\kappa > 0$, there exists β_{κ_1} such that, for all $\boldsymbol{\mu}' : \|\boldsymbol{\mu}-\boldsymbol{\mu}'\|_\infty \leq \beta_{\kappa_1}$:

$$\begin{aligned}
&\left|D(\boldsymbol{\mu}) - D(\boldsymbol{\mu}')\right| \leq \kappa, \\
&\left|\max_{x\in\mathcal{X}_F(\boldsymbol{\mu})} D(\boldsymbol{\mu},\boldsymbol{\omega},\neg x) - \max_{x\in\mathcal{X}_F(\boldsymbol{\mu}')} D(\boldsymbol{\mu}',\boldsymbol{\omega},\neg x)\right| \leq \kappa \quad \forall \boldsymbol{\omega}\in\Delta_K.
\end{aligned}$$

Furthermore, from Lemma 30, we know that, for all ϵ, there exists T_ϵ such that, for all $t \geq T_\epsilon$, on $\mathcal{E}(t)$, $\|\boldsymbol{\mu}-\boldsymbol{\mu}'\| \leq \epsilon$ holds for all $s \geq h(t)$ and $\boldsymbol{\mu}' \in C_s$. Thus, picking $\epsilon = \beta_{\kappa_1}$, we obtain that there exists $\bar{T}_1$ such that, for all $t \geq \bar{T}_1$, the first claim holds.

Second, from Corollary 1 and the compactness of $\mathcal{X}$ (*i.e.*, Assumption 1), we know that, for all $\kappa > 0$, there exists β_{κ_2} such that, for all $\boldsymbol{\mu}'$ such that $\|\boldsymbol{\mu}-\boldsymbol{\mu}'\|_\infty \leq \beta_{\kappa_2}$ and for all $x, x' \in \mathcal{X}$ such that $\|x-x'\|_\infty \leq \beta_{\kappa_2}$, we have that:

$$\left|D(\boldsymbol{\mu},\boldsymbol{\omega},\neg x) - D(\boldsymbol{\mu}',\boldsymbol{\omega},\neg x')\right| \leq \kappa, \quad \forall \boldsymbol{\omega}\in\Delta_K.$$

[22] We recall that sub-gaussianity implies $d(p,q) \geq \frac{(p-q)^2}{2\sigma^2}$.

From Definition 1, for all $\tilde{\epsilon} > 0$, there exists $T_{\tilde{\epsilon}}$ such that, for all $t \geq T_{\tilde{\epsilon}}$, on $\mathcal{E}(t)$, there exists $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$ such that $\|\bar{x} - x_s\| \leq \epsilon$ holds for all $s \geq h(t)$. Furthermore, from Lemma 30, we know that, for all ϵ, there exists T_ϵ such that, for all $t \geq T_\epsilon$, on $\mathcal{E}(t)$, $\|\boldsymbol{\mu} - \boldsymbol{\mu}'\| \leq \epsilon$ holds for all $s \geq h(t)$ and $\boldsymbol{\mu}' \in C_s$. Picking $\tilde{\epsilon} \leq \beta_{\kappa_2}$ and $\epsilon \leq \beta_{\kappa_2}$, we have that there exists $\bar{T}_2$ such that for all $t \geq \bar{T}_2$, the second claim holds. Taking $T_\kappa = \max\{\bar{T}_1, \bar{T}_2\}$ concludes the proof. □

Next, the following Lemma analyzes the behavior of the algorithm in terms of the l.h.s. of the stopping rule in Equation (13). Specifically, it relates the stopping rule to the characteristic time $T^*(\boldsymbol{\mu})$.

Lemma 32. *Suppose that Definition 1 holds. Then, for all $\kappa > 0$, there exists $T_\kappa \in \mathbb{N}$ such that, for all $t \geq T_\kappa$, on $\mathcal{E}(t)$, it holds that:*

$$\frac{1}{t} \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) \geq \frac{t - \lceil h(t) \rceil}{t} T^*(\boldsymbol{\mu})^{-1} - 3\kappa - \frac{K(1+\sqrt{t})}{t} C_{\boldsymbol{\mu}}$$

for some problem dependent constant $C_{\boldsymbol{\mu}} > 0$.

Proof. Let $\kappa > 0$ and consider $t \geq T_\kappa$ such that the results of Lemma 31 holds.

Then, we have that:

$$\frac{1}{t} \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) = \frac{1}{t} \max_{i \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{N}(t), \neg x) - h_1(t),$$

where $h_1(t)$ is given by:

$$h_1(t) := \frac{1}{t} \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{N}(t), \neg x) - \frac{1}{t} \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x).$$

From Lemma 31, we obtain $h_1(t) \leq \kappa$, thus leading to:

$$\frac{1}{t} \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) \geq \frac{1}{t} \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{N}(t), \neg x) - \kappa.$$

Furthermore, for all $\bar{x} \in \mathcal{X}^\star(\boldsymbol{\mu})$, we have that:

$$\begin{aligned}
\frac{1}{t} \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{N}(t), \neg x) &\geq \frac{1}{t} D(\boldsymbol{\mu}, \boldsymbol{N}(t), \neg \bar{x}) \\
&\geq \frac{1}{t} D\left(\boldsymbol{\mu}, \sum_{s=1}^{t} \boldsymbol{\omega}(s), \neg \bar{x}\right) - h_2(t),
\end{aligned}$$

where $h_2(t)$ is given by:

$$\begin{aligned}
h_2(t) &:= \frac{1}{t} \sup_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} \inf_{\boldsymbol{\lambda} \in \neg x} \sum_{k \in [K]} \left(\sum_{s=1}^{t} \omega_k(s) - N_k(t) \right) d(\mu_k, \lambda_k) \\
&\leq \frac{K(1+\sqrt{t})}{t} \sup_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} \inf_{\boldsymbol{\lambda} \in \neg x} \sum_{k \in [K]} d(\mu_k, \lambda_k) \\
&:= \frac{K(1+\sqrt{t})}{t} C_{\boldsymbol{\mu}},
\end{aligned}$$

where, in the first inequality, we used Lemma 41 and in the last one the fact that the exponential family is regular and bounded. Thus, we obtained that, for all $\bar{x} \in \mathcal{X}^\star(\boldsymbol{\mu})$:

$$\frac{1}{t} \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) \geq \frac{1}{t} D\left(\boldsymbol{\mu}, \sum_{s=1}^{t} \boldsymbol{\omega}(s), \neg \bar{x}\right) - \kappa - \frac{K(1+\sqrt{t})}{t} C_{\boldsymbol{\mu}}.$$

Therefore, it also holds for $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$ such that, on $\mathcal{E}(t)$,

$$\left| D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg\bar{x}) - D(\boldsymbol{\mu}', \boldsymbol{\omega}, \neg x_s) \right| \leq \kappa, \quad \forall \boldsymbol{\omega} \in \Delta_K, \boldsymbol{\mu}' \in C_s, s \geq h(t),$$

where the existence of such an $\bar{x}$ is given in Lemma 31. Thus, focus on $\frac{1}{t} D\left(\boldsymbol{\mu}, \sum_{s=1}^t \boldsymbol{\omega}(s), \neg\bar{x}\right)$:

$$\begin{aligned}
\frac{1}{t} D\left(\boldsymbol{\mu}, \sum_{s=1}^t \boldsymbol{\omega}(s), \neg\bar{x}\right) &\geq \frac{1}{t} \sum_{s=h(t)}^t D(\boldsymbol{\mu}, \boldsymbol{\omega}(s), \neg\bar{x}) \\
&= \frac{1}{t} \sum_{s=h(t)}^t D(\boldsymbol{\mu}'(s), \boldsymbol{\omega}(s), \neg x_s) - h_3(t) \\
&= \frac{1}{t} \sum_{s=h(t)}^t D(\boldsymbol{\mu}'(s)),
\end{aligned}$$

where $\boldsymbol{\mu}'(s) \in C_s$ is such that $\boldsymbol{\omega}(s) \in \boldsymbol{\omega}^*(\boldsymbol{\mu}'(s), \neg x_s)$ and $h_3(t)$ is given by:

$$h_3(t) = \frac{1}{t} \sum_{s=h(t)}^t \left(D(\boldsymbol{\mu}'(s), \boldsymbol{\omega}(s), \neg x_s) - D(\boldsymbol{\mu}, \boldsymbol{\omega}(s), \neg\bar{x}) \right) \leq \kappa,$$

where in the inequality step we used Lemma 31. Finally, we have that:

$$\frac{1}{t} \sum_{s=h(t)}^t D(\boldsymbol{\mu}'(s)) = \frac{(t - h(t))}{t} T^*(\boldsymbol{\mu})^{-1} - h_4(t),$$

where $h_4(t)$ is given by:

$$h_4(t) = \frac{1}{t} \sum_{s=h(t)}^t D(\boldsymbol{\mu}) - D(\boldsymbol{\mu}'(s)) \leq \kappa,$$

where the inequality step follows from Lemma 31, thus concluding the proof. □

Finally, Lemma 32 allows to prove the optimality of Sticky Sequence Track-and-Stop.

Theorem 3. *Sticky Sequence Track-and-Stop, equipped with a convergent selection rule, is δ-correct and asymptotically optimal,* i.e., $\limsup_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} \leq T^*(\boldsymbol{\mu})$.

Proof. Let $\kappa > 0$. From Lemma 32, for $t \geq T_\kappa$, on $\mathcal{E}(t)$, it holds that:

$$\frac{1}{t} \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) \geq \frac{t - \lceil h(t) \rceil}{t} T^*(\boldsymbol{\mu})^{-1} - 3\kappa - \frac{K(1+\sqrt{t})}{t} C_{\boldsymbol{\mu}}, \tag{16}$$

where $C_{\boldsymbol{\mu}}$ is a problem-dependent constant.

Let $\gamma \in \left(0, \frac{T^*(\boldsymbol{\mu})^{-1}}{4}\right]$, take $\kappa = \kappa(\gamma) \leq \frac{\gamma}{12}$. Furthermore, consider T_γ as the smallest integer $n \in \mathbb{N}$ such that $\frac{h(n)}{n} T^*(\boldsymbol{\mu})^{-1} \leq \frac{\gamma}{4}$ and $\frac{K(1+\sqrt{n})}{n} C_{\boldsymbol{\mu}} \leq \frac{\gamma}{4}$. Then, for $t \geq T_\gamma + T_{\kappa(\gamma)}$, Equation (16) implies:

$$\frac{1}{t} \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) \geq T^*(\boldsymbol{\mu})^{-1} - \gamma.$$

Applying Lemma 42 with $\alpha = L = \gamma$ and $D = T^*(\boldsymbol{\mu})^{-1}$, we have that for $t \geq \lceil T_0(\gamma, \gamma, \delta, T^*(\boldsymbol{\mu})^{-1}) + T_\gamma + T_{\kappa(\gamma)} \rceil$, we have that:

$$\frac{1}{t} \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) \geq \frac{\beta_{t,\delta}}{t},$$

which implies that, on the good event, the algorithm stops using at most $\lceil T_0(\gamma, \gamma, \delta, T^*(\boldsymbol{\mu})^{-1}) + T_\gamma + T_{\kappa(\gamma)} \rceil$ samples. Applying Lemma 29 we, thus, obtain:

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta] \leq T_0(\gamma, \gamma, \delta, T^*(\boldsymbol{\mu})^{-1}) + T_\gamma + T_{\kappa(\gamma)} + \sum_{s=0}^{t} \mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}(s)^c).$$

From Lemma 14 in Degenne and Koolen [2019], $\sum_{s=0}^{t} \mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}(s)^c)$ is finite. Thus, using the expression of $T_0(\gamma, \gamma, \delta, T^*(\boldsymbol{\mu})^{-1})$ given by Lemma 42, we have that:

$$\limsup_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} \leq \frac{1}{T^*(\boldsymbol{\mu})^{-1} - 2\gamma}.$$

Letting $\gamma \to 0$ concludes the proof. □

E.3 Algorithms for Converging Sequences

E.3.1 Sticky Sequence Track-and-Stop for the case in which $|\mathcal{X}_F(\mu)|$ is unique

We now show that Definition 1 holds whenever $\mathcal{X}_F(\boldsymbol{\mu})$ is finite by simply picking any $x_t \in \mathcal{X}_t$.

Lemma 33. *Let $\{x_t\}_t$ be such that $x_t \in \mathcal{X}_t$ for all $t \in \mathbb{N}$. Then, if $|\mathcal{X}_F(\boldsymbol{\mu})| = 1$ for all $\boldsymbol{\mu} \in \Theta^K$, then Definition 1 holds, i.e., for all $\epsilon > 0$, there exists T_ϵ such that, for all $t \geq T_\epsilon$, on $\mathcal{E}(t)$, it holds that $|x_s - \bar{x}| \leq \epsilon$ for all $s \geq h(t)$ for $\bar{x} = \mathcal{X}_F(\boldsymbol{\mu})$.*

Proof. From Lemma 1, $\boldsymbol{\mu} \rightrightarrows \mathcal{X}_F(\boldsymbol{\mu})$ is upper hemicontinuous. Furthermore, since $\mathcal{X}_F(\boldsymbol{\mu})$ is single-valued, then $\boldsymbol{\mu} \rightrightarrows \mathcal{X}_F(\boldsymbol{\mu})$ is a continuous function. Thus, for all $\epsilon > 0$, there exists $\eta_\epsilon > 0$ such that, for all $\boldsymbol{\mu}' \in \mathcal{B}_{\eta_\epsilon}(\boldsymbol{\mu})$, we have that $\|\mathcal{X}_F(\boldsymbol{\mu}) - \mathcal{X}_F(\boldsymbol{\mu}')\|_\infty \leq \epsilon$.

Furthermore, from Lemma 30, for all $\kappa > 0$, there exists $T_\kappa \in \mathbb{N}$ such that, for all $t \geq T_\kappa$, on $\mathcal{E}_t$, $\|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \kappa$ for all $\boldsymbol{\mu}' \in C_s$ and $s \geq h(t)$. It then follows that, for $\kappa = \eta_\epsilon$, we have that $\|\mathcal{X}_F(\boldsymbol{\mu}) - \mathcal{X}_F(\boldsymbol{\mu}')\|_\infty \leq \epsilon$ for all $\boldsymbol{\mu}' \in C_s$ and $s \geq h(t)$. Then, since $x_s \in \mathcal{X}_s$ and $\mathcal{X}_s = \bigcup_{\boldsymbol{\mu}' \in C_s} \mathcal{X}_F(\boldsymbol{\mu})$, we also have that, on $\mathcal{E}(t)$, $\|x_s - \mathcal{X}_F(\boldsymbol{\mu})\|_\infty \leq \epsilon$, $\forall s \geq h(t)$, which concludes the proof. □

As for the case in which $|\mathcal{X}_F(\boldsymbol{\mu})| = 1$, we note that Lemma 33 is actually slightly stronger than Definition 1 as we know exactly the answer toward which we are converging.

E.3.2 Sticky Sequence Track-and-Stop for the case in which $\mathcal{X} \subseteq \mathbb{R}$

We show that Definition 1 holds when $\mathcal{X} \subseteq \mathbb{R}$ and for $x_t \in \operatorname{argmin}_{x \in \mathcal{X}_t} x$. In the following, we assume that x_t is attained within $\mathcal{X}_t$. If this is not the case, the inf will be for sure attained on the closure of $\mathcal{X}_t$ (as $\mathrm{cl}(\mathcal{X}_t)$ is a compact set). In this case, we can simply pick any $x_t \in \mathcal{X}_t$ arbitrary close to $\operatorname{argmin}_{x \in \mathrm{cl}(\mathcal{X}_t)} x$ and the proof follows by identical arguments.

Lemma 34. *Let $\mathcal{X} \subset \mathbb{R}$. and $\{x_t\}_t$ be such that $x_t \in \operatorname{argmin}_{x \in \mathcal{X}_t} x$ for all $t \in \mathbb{N}$. Then, Definition 1 holds,* i.e., *for all $\epsilon > 0$, there exists T_ϵ such that, for all $t \geq T_\epsilon$, on $\mathcal{E}(t)$, it holds that $|x_s - \bar{x}| \leq \epsilon$ for all $s \geq h(t)$ for $\bar{x} \in \operatorname{argmin}_{x \in \mathcal{X}_F(\boldsymbol{\mu})} x$.*

Proof. By upper hemicontinuity of $\mathcal{X}_F(\boldsymbol{\mu})$, we have that, for all $\epsilon > 0$, there exists $\rho_\epsilon > 0$ such that, if $\|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \rho_\epsilon$, then $\mathcal{X}_F(\boldsymbol{\mu}') \subset \mathcal{B}_\epsilon(\mathcal{X}_F(\boldsymbol{\mu}))$.

Let T_ϵ be such that, for all $t \geq T_\epsilon$, under $\mathcal{E}(t)$, it holds that $\|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \rho_\epsilon$ for all $\boldsymbol{\mu}' \in C_s$ and $s \geq h(t)$. From Lemma 30 we are guaranteed that such T_ϵ exists.

Then, it follows that, on $|x_s - \bar{x}| \leq \epsilon$ for all $s \geq h(t)$ on the good event $\mathcal{E}(t)$. Indeed, (i) $x_s \leq \bar{x}$ since $\boldsymbol{\mu} \in C_s$ on the good event, (ii) $x_s \geq \bar{x} - \epsilon$ since $\mathcal{X}_F(\boldsymbol{\mu}') \in \mathcal{B}_\epsilon(\mathcal{X}_F(\boldsymbol{\mu}))$ and $\bar{x} \in \operatorname{argmin}_{x \in \mathcal{X}_F(\boldsymbol{\mu})} x$. □

As a minor remark, we observe that Lemma 34 is actually slightly stronger than Definition 1 as we know exactly the answer toward which we are converging.

Algorithm 2 A general procedure for selecting answers.

Require: Radius function $\rho : \mathbb{N} \to \mathbb{R}$, answer space $\mathcal{X}$
Initialization: Pick any $\bar{x}_1 \in \mathcal{P}_1 : \mathcal{B}_{\rho_1}(\bar{x}_1) \cap \mathcal{X}_1 \neq \emptyset$. Set $x_1 \in \mathcal{B}_{\rho_1}(\bar{x}_1) \cap \mathcal{X}_1$. Set $\mathcal{H}_1 = \{(\bar{x}_1, \rho_1)\}$
Selection Rule at step t:
Let $\mathcal{S}_t = \{(x, \rho) \in \mathcal{H}_{t-1} : \forall (\bar{x}_s, \rho_s) \in \mathcal{H}_{t-1},\ \rho_s > \rho \text{ and } \mathcal{B}_{\rho_s}(\bar{x}_s) \cap \mathcal{X}_t \neq \emptyset\}$
Let $\bar{s}$ be the time corresponding to the element $(\bar{x}_s, \rho_s) \in \mathcal{S}_t$ with the smallest radius ρ_s
Let $\overline{\mathcal{H}}_t = \emptyset$
for $s = \bar{s} + 1, \dots, t$ **do**
  Pick any $\bar{x}_s \in \mathcal{P}_s$ such that $\mathcal{B}_{\rho_s}(\bar{x}_s) \cap \mathcal{B}_{\rho_{s-1}}(\bar{x}_{s-1}) \cap \mathcal{X}_t \neq \emptyset$
  $\overline{\mathcal{H}}_t = \overline{\mathcal{H}}_t \cup (\bar{x}_s, \rho_s)$
end for
Pick any $x_t \in \mathcal{B}_{\rho_t}(\bar{x}_t) \cap \mathcal{X}_t$
Update History $\mathcal{H}_t = \{(x, \rho) \in \mathcal{H}_{t-1} : \rho \geq \rho_{\bar{s}}\} \cup \overline{\mathcal{H}}_t$

E.3.3 Sticky Sequence Track-and-Stop for the case in which $|\mathcal{X}_F(\mu)|$ is finite

We show that Definition 1 holds when $|\mathcal{X}_F(\boldsymbol{\mu})|$ is finite and $x_t \in \operatorname{argmin}_{x \in \mathcal{X}_t} \|x - x_{t-1}\|_\infty$. As above, we assume that x_t is attained within $\mathcal{X}_t$. Again, if this is not the case, the inf will be for sure attained on the closure of $\mathcal{X}_t$ and we can simply pick any $x_t \in \mathcal{X}_t$ arbitrary close to $\operatorname{argmin}_{x \in \mathrm{cl}(\mathcal{X}_t)} x$. The proof follows by identical arguments.

Lemma 35. *Let $\{x_t\}_t$ be such that $x_t \in \operatorname{argmin}_{x \in \mathcal{X}_t} \|x - x_{t-1}\|_\infty$ for all $t \in \mathbb{N}$. Suppose that $|\mathcal{X}_F(\boldsymbol{\mu})| \leq M$ for all $\boldsymbol{\mu} \in \Theta^K$, then Definition 1 holds,* i.e., *for all $\epsilon > 0$, there exists T_ϵ, such that, for all $t \geq T_\epsilon$, under $\mathcal{E}(t)$, there exists $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$, such that $\|\bar{x} - x_s\|_\infty \leq \epsilon$ for all $s \geq h(t)$.*

Proof. Let ρ be such that $\|x_1 - x_2\|_\infty > \rho$ for all $x_1, x_2 \in \mathcal{X}_F(\boldsymbol{\mu})$, $x_1 \neq x_2$. Then, since $\mathcal{X}_F(\boldsymbol{\mu})$ is upper hemicontinuous (Lemma 1), there exists $\eta_\rho > 0$ such that, for all $\boldsymbol{\mu}' : \|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \eta_\rho \implies \mathcal{X}_F(\boldsymbol{\mu}') \subset \mathcal{B}_\rho(\mathcal{X}_F(\boldsymbol{\mu}))$. Since $|\mathcal{X}_F(\boldsymbol{\mu})| < M$, it follows that:

$$\bigcup_{\boldsymbol{\mu} : \|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \eta_\rho} \mathcal{X}_F(\boldsymbol{\mu}') \subseteq \bigcup_{x \in \mathcal{X}_F(\boldsymbol{\mu})} \mathcal{B}_\rho(x),$$

and, moreover, $\mathcal{B}_\rho(x_1) \cap \mathcal{B}_\rho(x_2) = \emptyset$ for all $x_1, x_2 \in \mathcal{X}_F(\boldsymbol{\mu})$, $x_1 \neq x_2$.

Let $t \geq \widetilde{T}_{\eta_\rho}$, where $\widetilde{T}_{\eta_\rho}$ is such that, for all $t \geq \widetilde{T}_{\eta_\rho}$, on $\mathcal{E}(t)$, $\|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \eta_\rho$ for all $\boldsymbol{\mu}' \in C_s$ and all $s \geq h(t)$. Such $\widetilde{T}_{\eta_\rho}$ is guaranteed to exists due to Lemma 30. Then, on $\mathcal{E}(t)$, $\|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \eta_\rho$ for all $\boldsymbol{\mu}' \in C_s$ and all $s \geq h(t)$.

Now, we observe that, for $t \geq \widetilde{T}_{\eta_\rho}$ and $s \geq h(t)$, on $\mathcal{E}(t)$, $\|x_s - \bar{x}\|_\infty \leq \rho$ for some $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$, and, furthermore, $\|x_s - x\|_\infty > \rho$ for all $x \in \mathcal{X}_F(\boldsymbol{\mu}) \setminus \bar{x}$. This is due to the fact that $\boldsymbol{\mu} \in C_s$ for all $s \geq h(t)$ (and, consequently, $\mathcal{X}_F(\boldsymbol{\mu}) \in \mathcal{X}_s$), $\|x_1 - x_2\|_\infty > \rho$ for all $x_1, x_2 \in \mathcal{X}_F(\boldsymbol{\mu})$, and thanks to the fact that the selection rule of x_s is such that it always select the feasible solution which is closest to the previous point.

Thus, Definition 1 directly follows by considering any $t \geq T_\epsilon \coloneqq \max\{\widetilde{T}_\epsilon, \widetilde{T}_\rho\}$. Indeed, by the reasoning above, we have that there exists $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$ such that $\|x_s - \bar{x}\| \leq \min\{\rho, \epsilon\} \leq \epsilon$ for all $s \geq h(t)$ under the good event $\mathcal{E}(t)$.

□

E.3.4 Sticky Sequence Track and Stop with Adaptive Discretization

We now show that there exists a general algorithm for ensuring Definition 1 in arbitrary compact spaces. The pseudocode of the selection rule for the answer x_t can be found in Algorithm 2. Before detailing the algorithm we introduce some notation.

Preliminary definitions We consider a radius function $\rho : \mathbb{N} \to \mathbb{R}_{>0}$ such that $\rho(t+1) < \rho(t)$ and $\lim_{t \to \infty} \rho(t) = 0$. Specifically, we choose $\rho(t) = 2^{-t}$. In the following, with some abuse of notation, we write $\rho(t) = \rho_t$. Note that, for any ρ_t, the corresponding time t is uniquely identified by ρ_t due to the fact that $\rho(t)$ is invertible.

For the answer set $\mathcal{X}$ and any $t \in \mathbb{N}$, we write write $\mathcal{P}_t = \{x_i\}_{i=1}^{n_{\rho_t}}$ (or, with some abuse of notation, $\mathcal{P}_{\rho_t}$) to denote the centers of a cover of $\mathcal{X}$ that uses balls of radius ρ_t centered in points $x_i \in \mathcal{P}_t$.

Explanation of the algorithm Algorithm 2 works by combining the progressive discretization of $\mathcal{X}$ together with a "history mechanism".

In the first turn, the algorithm selects any point $\bar{x}_1 \in \mathcal{P}_1$ such that $\mathcal{B}_{\rho_1}(\bar{x}_1) \cap \mathcal{X}_1 \neq \emptyset$ and it picks any answer $x_1 \in \mathcal{B}_{\rho_1}(\bar{x}_1) \cap \mathcal{X}_1$. Then, it initializes a "history" $\mathcal{H}_1 = \{(\bar{x}_1, \rho_1)\}$.

Assume now that a history $\mathcal{H}_{t-1}$ of $t-1$ tuples $\{(\bar{x}_s, \rho_s)\}_{s=1}^{t-1}$ is given to the algorithm (we will explain shortly how $\mathcal{H}_t$ is defined and updated by the algorithm). The algorithm decides the answer x_t to play as follows. Among all elements of $\mathcal{H}_{t-1}$, the algorithm first selects an index $\bar{s}$, which is the one with the smallest radius (*i.e.*, the greatest index $\bar{s} \leq t-1$) that guarantees that the intersection of $\mathcal{B}_{\rho_s}(\bar{x}_s)$ and $\mathcal{X}_t$ is non-empty for all $s \leq \bar{s}$ (Line 3 and Line 4). Intuitively, the region $\mathcal{B}_{\rho_{\bar{s}}}(\bar{x}_{\bar{s}})$ can be seen as an anchoring mechanism that constrains the search for a good answer in $\mathcal{X}_F(\boldsymbol{\mu})$ towards previously selected points (for further explanation, see the remark of $\bar{s}$ below). When such $(\bar{x}_{\bar{s}}, \rho_{\bar{s}})$ is selected, then the tuples (x_s, ρ_s) with $s > \bar{s}$ are discarded from the history set, and $\mathcal{H}_t$ is "repopulated" by picking some $(\bar{x}_s, \rho_s)$ such that $\bar{x}_s$ is in the centers of the cover at time s and such that $\mathcal{B}_{\rho_s}(\bar{x}_s) \cap \mathcal{B}_{\rho_{s-1}}(\bar{x}_{s-1}) \cap \mathcal{X}_t \neq \emptyset$ (Line 7). In Algorithm 2 the "repopulation" of $\mathcal{H}_t$ is formalized trough $\overline{\mathcal{H}}_t$ (Line 8 and Line 11).

Finally, Algorithm 2 simply selects any $x_t \in \mathcal{B}_{\rho_t}(\bar{x}_t) \cap \mathcal{X}_t$ (Line 10).

Remark on $\bar{s}$ We observe that, for all $s > \bar{s}$, there exists an element in position $s' \in \{s+1, s\}$ such that $\mathcal{B}_{\rho_{s'}}(\bar{x}_{s'}) \cap \mathcal{X}_t = \emptyset$. We recall that:

- $\mathcal{X}_t$ represents the set of candidate answers that correspond to models within a confidence region around $\hat{\boldsymbol{\mu}}_t$;
- Algorithm 2 keeps in $\mathcal{H}_t$ only elements of $\mathcal{H}_{t-1}$ whose index is less than $\bar{s}$, *i.e.*, for all $s \leq \bar{s}$, it holds that $\mathcal{B}_{\rho_s}(\bar{x}_s) \cap \mathcal{X}_t \neq \emptyset$.

These two facts can be interpreted as a *backtracking* operation that is needed to guide the search of the algorithm toward some $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$. Indeed, for all $s > \bar{s}$, there is a point within the sequence $\{\bar{s}, \dots, s\}$ whose ball does not intersect the set of candidate answers $\mathcal{X}_t$ (which, on the good event, contains $\mathcal{X}_F(\boldsymbol{\mu})$).

Remark on notation In the rest of this section, for the sake of clarity, we will use the following convention:

- $\bar{x} \in \mathcal{X}$ denotes the answer in $\mathcal{X}_F(\boldsymbol{\mu})$ towards which Algorithm 2 is converging (or, whenever needed, answers in $\mathcal{X}_F(\boldsymbol{\mu})$);
- x_t denotes the answer selected by Algorithm 2 at step t;
- $\bar{x}_s^{(t)}$ denotes the center of the cover of radius ρ_s (*i.e.*, $\bar{x}_s \in \mathcal{P}_s$) within the history set at step t, *i.e.*, $(\bar{x}_s^{(t)}, \rho_s) \in \mathcal{H}_t$. Note that here, we are expanding the notation introduced above since the elements within $\mathcal{H}_t$ can change throughout the execution of the algorithm. This is to avoid potential ambiguities.

Given this convention, we observe that for all $t \in \mathbb{N}$, the selection rule of Algorithm 2 can be rewritten as $x_t \in \mathcal{B}_{\rho_t}(\bar{x}_t^{(t)})$ for some $\bar{x}_t^{(t)} \in \mathcal{P}_{\rho_t}$. Similarly, the history set $\mathcal{H}_t$ can be rewritten as $\mathcal{H}_t = \{(x_s^{(t)}, \rho_s)\}_{s=1}^{t}$. Observe that, due to the backtracking operation highlighted above, in principle it can happen that $x_s^{(t)} \neq x_s^{(t+1)}$.

Proof of convergence In the following, we prove that Algorithm 2 generates a converging sequence. We first give a proof outline and then we dive into the formal arguments.

The main idea is showing that for any $n \in \mathbb{N}$, and for sufficiently large t, under the good event $\mathcal{E}(t)$, there exists an element $(\bar{x}_n, \rho_n)$ that remains in $\mathcal{H}_s$ for all $s \geq h(t)$, *i.e.*, $\bar{x}_n = \bar{x}_n^{(s)}$ for all $s \geq h(t)$. This first result (which is formally stated in Lemma 37) intuitively says that, when enough information has been collected, then the algorithm is able to fix a region of arbitrary radius within its history set $\mathcal{H}_s$. As Lemma 38 then shows, this is enough to ensure that the algorithm is converging to some $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$. Indeed, we will show that the

aforementioned $\bar{x}_n$ is close to some $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$, and, by the design of the algorithm (*i.e.*, Line 8), the answer selected at step s (*i.e.*, x_s) will be close to $\bar{x}_n$ for all $s \geq h(t)$. By a triangular inequality argument, this implies that x_s remains close to some $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$, thus leading to the desired convergence property.

Now, before moving to Lemma 37, we first prove a technical result which will be used within the proof of Lemma 37.

Lemma 36. *Consider $\boldsymbol{\mu} \in \mathcal{M}$, $n \in \mathbb{N}$ and $\rho = \rho_n$. Suppose that there exists $x \in \mathcal{P}_\rho$ such that $\mathcal{B}_\rho(x) \cap \mathcal{X}_F(\boldsymbol{\mu}) = \emptyset$. Then, it holds that:*

$$\eta_\rho = \min_{\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})} \min_{\substack{x \in \mathcal{P}_\rho: \\ \mathcal{B}_\rho(x) \cap \mathcal{X}_F(\boldsymbol{\mu}) = \emptyset}} \min_{x' \in \mathcal{B}_\rho(x)} \|x' - \bar{x}\|_\infty > 0.$$

where η_ρ is well defined and $\eta_\rho > 0$.

Proof. Consider

$$\eta_\rho := \inf_{\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})} \inf_{\substack{x \in \mathcal{P}_\rho: \\ \mathcal{B}_\rho(x) \cap \mathcal{X}_F(\boldsymbol{\mu}) = \emptyset}} \inf_{x' \in \mathcal{B}_\rho(x)} \|x' - \bar{x}\|_\infty.$$

Then, fix any $x \in \mathcal{P}_\rho$ and $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$ and let $\phi(\bar{x}, x) = \inf_{x' \in \mathcal{B}_\rho(x)} \|x' - \bar{x}\|_\infty$, where the inf can be replaced by min (indeed, $\mathcal{B}_\rho(x)$ is compact and the inner infinite-norm function is continuous).

Now, let $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$ and consider:

$$\inf_{\substack{x \in \mathcal{P}_\rho: \\ \mathcal{B}_\rho(x) \cap \mathcal{X}_F(\boldsymbol{\mu}) = \emptyset}} \phi(\bar{x}, x).$$

We observe that, by assumption, the set over which we are optimizing is non-empty. Furthermore, it holds that:

$$\inf_{\substack{x \in \mathcal{P}_\rho: \\ \mathcal{B}_\rho(x) \cap \mathcal{X}_F(\boldsymbol{\mu}) = \emptyset}} \phi(\bar{x}, x) = \min_{\substack{x \in \mathcal{P}_\rho: \\ \mathcal{B}_\rho(x) \cap \mathcal{X}_F(\boldsymbol{\mu}) = \emptyset}} \phi(\bar{x}, x) := g(\bar{x}).$$

Indeed, the cardinality of the set $x \in \mathcal{P}_\rho : \mathcal{B}_\rho(x) \cap \mathcal{X}_F(\boldsymbol{\mu}) = \emptyset$ is finite since $\mathcal{P}_\rho$ is finite. Furthermore, $g(\bar{x})$ is continuous in $\bar{x}$ since it is a minimum over a finite set of continuous functions.

At this point, we note that the outer inf can be replaced by a min since that objective function is continuous in $\bar{x}$ and $\mathcal{X}_F(\boldsymbol{\mu})$ is compact. This shows that η_ρ is well defined.

It remains to prove that the value of the optimization problem is strictly greater than 0. Consider a triplet $(\bar{x}, x, x')$ that attains all the minimums of the optimization problem. Then, we need to prove that $\|\bar{x} - x'\|_\infty > 0$, i.e., that $\bar{x} \neq x'$. However, $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$ and $x' \notin \mathcal{X}_F(\boldsymbol{\mu})$ (since $x' \in \mathcal{B}_\rho(x)$ and $\mathcal{B}_\rho(x) \cap \mathcal{X}_F(\boldsymbol{\mu}) = \emptyset$), thus concluding the proof. □

We are now ready to show that, under the good event, the algorithm is able to identify good regions in which it will conduct the search.

Lemma 37. *For all $n \in \mathbb{N}$, there exists $T_n \in \mathbb{N}$ such that, for all $t \geq T_n$, under $\mathcal{E}(t)$, there exists $\bar{x}_n \in \mathcal{P}_n$ and:*

$$(\bar{x}_n, \rho_n) \in \mathcal{H}_s, \quad \forall s \geq h(t), \tag{17}$$

i.e., *for all $s \geq h(t)$, $\bar{x}_n = \bar{x}_n^{(s)}$. Furthermore, under $\mathcal{E}(t)$, there exists $\bar{x} \in \mathcal{X}$ such that*

$$\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu}) \cap \mathcal{B}_{\rho_n}(\bar{x}_n). \tag{18}$$

Proof. We start with some preliminary definitions.

Let $n \in \mathbb{N}$, and let us define the following set:

$$\mathbb{N}(n) = \{j \in \mathbb{N} : j \leq n \text{ and } \exists x \in \mathcal{P}_{\rho_j} : \mathcal{B}_{\rho_j}(x) \cap \mathcal{X}_F(\boldsymbol{\mu}) = \emptyset\}. \tag{19}$$

The set $\mathbb{N}(n)$ represents the subset of $[n]$ for which we can apply Lemma 36. Furthermore, let us define r_n as follows:

$$r_n = \frac{1}{2} \min_{j \in \mathbb{N}(n)} \eta_{\rho_j},$$

where η_{ρ_j} is as in Lemma 36.

Next, we recall that, by Lemma 30 and the upper hemicontinuity of $\mathcal{X}_F$ (*i.e.*, Lemma 1), for all $\Delta > 0$, there exists $\widetilde{T}_\Delta$ such that, for all $t \geq \widetilde{T}_\Delta$, under $\mathcal{E}(t)$, $\mathcal{X}_s \subseteq \mathcal{B}_\Delta(\mathcal{X}_F(\boldsymbol{\mu}))$ for all $s \geq h(t)$.

In the following, we will show that picking $\Delta = r_n$ leads to the desired result. Specifically, we will consider $T_n = \max\{\widetilde{T}_{r_n}, \bar{T}_n\})$, where $\bar{T}_n$ is such that $\rho_{h(\bar{T}_n)} \leq \rho_n$. Here, the requirement that $t \geq \bar{T}_n$ is a technicality that ensures that elements of the form $(\cdot, \rho_n)$ are already within the history set $\mathcal{H}_{h(t)}$. Indeed, at step $h(t)$, the history set $\mathcal{H}_{h(t)}$ only contains elements of the form $(\cdot, \rho)$ for $\rho \geq \rho_{h(t)}$. Intuitively, to ensure that $(\bar{x}_n^{(h(t)}, \rho_n) \in \mathcal{H}_{h(t)}$ for some $\bar{x}_n^{(h(t)} \in \mathcal{P}_{h(t)}$ we need $t \geq \bar{T}_n$. We now proceed with the crucial part of the proof, that is exploiting $t \geq \widetilde{T}_{r_n}$.

Observe that, to prove Equation (17), we need to show that there exists an element $(\bar{x}_n, \rho_n)$ that will not be removed from $\mathcal{H}_s$ in any $s \geq h(t)$, *i.e.*, $\exists \bar{x}_n : \bar{x}_n = \bar{x}_n^{(s)}$ for all $s \geq h(t)$. As discussed above, for $t \geq \bar{T}_n$, there is an element of the form $(\bar{x}_n^{(h(t))}, \rho_n)$ within $\mathcal{H}_{h(t)}$. In the following, we prove that such an element will not be eliminated from $\mathcal{H}_s$ as the execution proceeds.

By definition of Algorithm 2, an element $(\bar{x}_n^{(s)}, \rho_n)$ can fail to belong to $\mathcal{H}_{s+1}$ if and only if there exists $(\bar{x}_j^{(s)}, \rho_j) \in \mathcal{H}_s$ such that $j \leq n$ and $\mathcal{B}_{\rho_j}(\bar{x}_j^{(s)}) \cap \mathcal{X}_{s+1} = \emptyset$.

We proceed by contradiction. Suppose that there exists a $(\bar{x}_j^{(s)}, \rho_j) \in \mathcal{H}_s$ for which $\mathcal{B}_{\rho_j}(\bar{x}_j^{(s)}) \cap \mathcal{X}_{s+1} = \emptyset$ for some $j \leq n$. In the following, we will refer to the index j as the index of the element that *triggers* the elimination of $(\bar{x}_n^{(s)}, \rho_n)$ from $\mathcal{H}_{s+1}$. Now, we proceed by cases.

Case one: $j \notin \mathbb{N}(n)$ If $j \notin \mathbb{N}(n)$, then, we know by Equation (19) that $\mathcal{B}_{\rho_j}(\bar{x}_j^{(s)}) \cap \mathcal{X}_F(\boldsymbol{\mu}) \neq \emptyset$. Hence, since on the good event $\mathcal{X}_F(\boldsymbol{\mu}) \subseteq \mathcal{X}_s$ for all $s \geq h(t)$,[23] we also have that $\mathcal{B}_{\rho_j}(\bar{x}_j^{(s)}) \cap \mathcal{X}_s \neq \emptyset$, and, as a consequence, the elimination condition of $(\bar{x}_n^{(s)}, \rho_n)$ from $\mathcal{H}_{s+1}$ cannot be triggered by the element in position j, thus leading to the contradiction.

Case two: $j \in \mathbb{N}(n)$ If $j \in \mathbb{N}(n)$, instead, we prove that for $t \geq T_n$, $\mathcal{B}_{\rho_j}(\bar{x}_j^{(s)}) \cap \mathcal{X}_F(\boldsymbol{\mu}) \neq \emptyset$ and (as we have shown in the previous case), this in turn implies that the elimination condition of $(\bar{x}_n^{(s)}, \rho_n)$ from $\mathcal{H}_{s+1}$ cannot be triggered by the element in position j. To this end, the following argument shows that, for $t \geq \widetilde{T}_{r_n}$, all the elements $x \in \mathcal{P}_{\rho_j}$ such that $\mathcal{B}_{\rho_j}(x) \cap \mathcal{X}_F(\boldsymbol{\mu}) = \emptyset$ cannot satisfy $\mathcal{B}_{\rho_j}(x) \cap \mathcal{X}_s \neq \emptyset$ at any $s \geq h(t)$, and therefore, they cannot belong to $\mathcal{H}_s$. Specifically, suppose that there exists $x \in \mathcal{P}_{\rho_j}$ such that the following holds for $s \geq h(t)$:

$$\mathcal{B}_{\rho_j}(x) \cap \mathcal{X}_F(\boldsymbol{\mu}) = \emptyset \tag{20}$$

$$\mathcal{B}_{\rho_j}(x) \cap \mathcal{X}_s \neq \emptyset. \tag{21}$$

[23] Recall that $\boldsymbol{\mu} \in C_s$ by definition of $\mathcal{E}(t)$.

Then, under $\mathcal{E}(t)$, for all $s \geq h(t)$, and any $y \in \mathcal{B}_{\rho_j}(x) \cap \mathcal{X}_s$ (such y exists due to Equation (21)), there exists $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$ such that $\|y - \bar{x}\|_\infty \leq r_n$ (by definition of $\widetilde{T}_{r_n}$ and $t \geq T_{r_n}$). However:

$$\begin{aligned}
\|y - \bar{x}\|_\infty &\leq r_n \\
&< \eta_{\rho_j} && (r_n = \tfrac{1}{2}\min_{i \in \mathbb{N}(n)} \eta_{\rho_i} \text{ and } j \in \mathbb{N}(n)) \\
&\leq \min_{\substack{x' \in \mathcal{P}_{\rho_j}: \\ \mathcal{B}_{\rho_j}(x') \cap \mathcal{X}_F(\boldsymbol{\mu}) = \emptyset}} \min_{x'' \in \mathcal{B}_{\rho_j}(x')} \|x'' - \bar{x}\|_\infty && (\text{Def. of } \eta_{\rho_j}) \\
&\leq \min_{x'' \in \mathcal{B}_{\rho_j}(x)} \|x'' - \bar{x}\|_\infty && (x \in \mathcal{P}_{\rho_j} : \mathcal{B}_{\rho_j}(x) \cap \mathcal{X}_F(\boldsymbol{\mu}) = \emptyset, \textit{ i.e.}, \text{ Equation (20)}) \\
&\leq \|y - \bar{x}\|_\infty. && (y \in \mathcal{B}_{\rho_j}(x))
\end{aligned}$$

Thus, we have shown that $\|y - \bar{x}\|_\infty < \|y - \bar{x}\|_\infty$ which leads to a contradiction. This concludes the proof of Equation (17) since we have shown that for all $s \geq h(t)$ all the conditions that can trigger the elimination of $(\bar{x}_n^{(s)}, \rho_n)$ from $\mathcal{H}_{s+1}$ cannot be triggered.

Now, concerning Equation (18), we observe that the proof of Equation (17) already used in its argument the existence, for all $j \leq n$, of $\mathcal{X}_F(\boldsymbol{\mu}) \cap \mathcal{B}_{\rho_j}(\bar{x}_j^{(s)}) \neq \emptyset$. Thus it holds also for $j = n$ and $\bar{x}_n^{(s)} = \bar{x}_n$ for all $s \geq h(t)$. □

We are now ready to prove that Definition 1 is satisfied for the sequence $\{x_t\}_t$ generated by Algorithm 2.

Lemma 38. *Consider that $\{x_t\}$ is given by Algorithm 2. Then, Definition 1 holds, i.e., for all $\epsilon > 0$, there exists $T_\epsilon \in \mathbb{N}$ such that, for all $t \geq T_\epsilon$, on $\mathcal{E}(t)$, there exists $\bar{x} \in \mathcal{X}_F(\boldsymbol{\mu})$ such that $\|\bar{x} - x_s\|_\infty \leq \epsilon$ for all $s \geq h(t)$.*

Proof. From Lemma 37 we know that, for all $n \in \mathbb{N}$, there exists $\widetilde{T}_n$ such that, for all $t \geq \widetilde{T}_n$, under $\mathcal{E}(t)$, there exists $\bar{x}_n \in \mathcal{P}_n$ and $(\bar{x}_n, \rho_n) \in \mathcal{H}_s, \forall s \geq h(t)$. Let $\epsilon > 0$ and consider $T_\epsilon = \widetilde{T}_{\bar{n}}$ where $\bar{n}$ is the smallest integer that verifies $\rho_{\bar{n}-1} \leq \epsilon/4$.

It follows that for all $t \geq T_\epsilon$ the following properties hold under $\mathcal{E}_t$:

$$\exists \bar{x}_{\bar{n}} \in \mathcal{P}_{\bar{n}} : (\bar{x}_{\bar{n}}, \rho_{\bar{n}}) \in \mathcal{H}_s \ \forall s \geq h(t), \text{ and } \rho_{\bar{n}} \leq \epsilon/4. \tag{22}$$

Observe that, as a direct consequence of Equation (22), we have that:

$$\rho_s \leq \rho_{\bar{n}} \quad \forall s \geq h(t). \tag{23}$$

Indeed, from Lemma 37, we have that $(\bar{x}_{\bar{n}}, \rho_{\bar{n}}) \in \mathcal{H}_s$ for all $s \geq h(t)$. Then, it follows by definition of Algorithm 2 that $\rho_{\bar{n}}$ is related to some step $\bar{n}$ which is at most $h(t)$ (since that element need to be in the history set already at step $h(t)$). In other words, $h(t) \geq \bar{n}$.

Furthermore, from Lemma 37 and $\bar{x}_{\bar{n}}$ as in Equation (22), we have that:

$$\exists \bar{x} \in \mathcal{X}_F(\boldsymbol{\mu}) \cap \bar{x} \in \mathcal{B}_{\rho_{\bar{n}}}(\bar{x}_{\bar{n}}). \tag{24}$$

Now, by definition of Algorithm 2, $x_s \in \mathcal{B}_{\rho_s}(\bar{x}_s^{(s)})$ for all $s \geq 1$ for some $\bar{x}_s^{(s)} \in \mathcal{P}_{\rho_s}$. Let $\bar{x}_{\bar{n}}$ be as in Equation (22) and $\bar{x}$ as in Equation (24). In the following, we will prove that under $\mathcal{E}(t)$, $\|x_s - \bar{x}\|_\infty \leq \epsilon$ for all $s \geq h(t)$, which is exactly Definition 1. Specifically, we have that:

$$\begin{aligned}
\|x_s - \bar{x}\|_\infty &\leq \|x_s - \bar{x}_{\bar{n}}\|_\infty + \|\bar{x}_{\bar{n}} - \bar{x}\|_\infty \\
&\leq \|x_s - \bar{x}_{\bar{n}}\|_\infty + \rho_{\bar{n}} && (\text{Equation (24)}) \\
&\leq \|x_s - \bar{x}_{\bar{n}}\|_\infty + \frac{\epsilon}{4} && (\text{Def. of } \bar{n}) \\
&\leq \|x_s - \bar{x}_s^{(s)}\|_\infty + \|\bar{x}_s^{(s)} - \bar{x}_{\bar{n}}\|_\infty + \frac{\epsilon}{4} \\
&\leq \rho_s + \|\bar{x}_s^{(s)} - \bar{x}_{\bar{n}}\|_\infty + \frac{\epsilon}{4} && (\text{Def. of } x_s) \\
&\leq \rho_{\bar{n}} + \|\bar{x}_s^{(s)} - \bar{x}_{\bar{n}}\|_\infty + \frac{\epsilon}{4} && (\text{Equation (23)}) \\
&\leq \|\bar{x}_s^{(s)} - \bar{x}_{\bar{n}}\|_\infty + \frac{\epsilon}{2} && (\text{Def. of } \bar{n})
\end{aligned}$$

At this point, it remains to upper bound $\|\bar{x}_s^{(s)} - \bar{x}_{\bar{n}}\|_\infty$, which we analyze with a telescoping argument. Recall that $\bar{n} \leq s$ for all $s \geq h(t)$. Then,

$$\|\bar{x}_s^{(s)} - \bar{x}_{\bar{n}}\|_\infty \leq \sum_{j=\bar{n}}^{s} \|\bar{x}_j^{(s)} - \bar{x}_{j+1}^{(s)}\|_\infty,$$

where we introduced all the elements in the history $\mathcal{H}_s$ from step $\bar{n}$ to s and we have used that, due to Lemma 37, $\bar{x}_{\bar{n}} = \bar{x}_{\bar{n}}^{(s)}$ for all $s \geq h(t)$. Then, by construction (*i.e.*, Line 7 in Algorithm 2), we have that $\mathcal{B}_{\rho_j}(\bar{x}_j^{(s)}) \cap \mathcal{B}_{\rho_{j+1}}(\bar{x}_{j+1}^{(s)}) \neq \emptyset$. Thus, $\|\bar{x}_j^{(s)} - \bar{x}_{j+1}^{(s)}\|_\infty \leq \rho_j + \rho_{j+1} \leq 2\rho_j$. It then follows that:

$$\|\bar{x}_s^{(s)} - \bar{x}_{\bar{n}}\|_\infty \leq 2\sum_{j=\bar{n}}^{s} \rho_j \leq 2\sum_{j=\bar{n}}^{+\infty} \frac{1}{2^j} \leq 2\frac{1}{2^{\bar{n}-1}} = 2\rho_{\bar{n}-1} \leq \frac{\epsilon}{2}.$$

Combining these results, we have obtained that $\|x_s - \bar{x}\|_\infty \leq \epsilon$ under $\mathcal{E}(t)$ for sufficiently large t. This concludes the proof. □

E.4 What happens when $\{x_t\}$ is not converging

In this section, we discuss what happens when the underlying sequence is not a converging one. Specifically, what kind of theoretical guarantees can we obtain? Note that answering this question also provides the theoretical guarantees of TaS and Sticky-TaS whenever they fail to generate a converging sequence. Indeed, as we discussed in Section 5.1, our framework actually generalizes both TaS and Sticky-TaS (according to the rule with which x_t is selected).

On a theoretical level, the main result of this section is Theorem 4. Specifically, Theorem 4 shows that, when the sequence x_t is not a converging one, we are only able to guarantee a results of the following form:

$$\limsup_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} \leq \max_{\boldsymbol{\omega} \in \mathrm{conv}(\boldsymbol{\omega}^\star(\boldsymbol{\mu}))} \min_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} \frac{1}{D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x)}. \tag{25}$$

The rest of this section is structured as follows. First, in Appendix E.4.1, we provide the formal statement and the proof of Equation (25). The formal result is given in Theorem 4. Then, in Appendix E.4.2, we provide, as curiosity, a theoretical example of an application for which we can quantify the gap between the result of Theorem 3 and Theorem 4.

E.4.1 Theoretical Results

To prove Theorem 4, we first show a result that combines Lemma 1 and the good event.

Lemma 39. *For all $\kappa > 0$, there exists $T_\kappa \in \mathbb{N}$ such that, for all $t \geq T_\kappa$, on $\mathcal{E}(t)$, it holds that, for all $s \geq h(t)$, and all $\boldsymbol{\mu}' \in C_s$:*

$$\Big| \max_{x \in \mathcal{X}_F(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x) - \max_{x \in \mathcal{X}_F(\boldsymbol{\mu}')} D(\boldsymbol{\mu}', \boldsymbol{\omega}, \neg x) \Big| \leq \kappa, \ \forall \boldsymbol{\omega} \in \Delta_K,$$
$$\inf_{\boldsymbol{\omega} \in \boldsymbol{\omega}^*(\boldsymbol{\mu})} \|\boldsymbol{\omega} - \boldsymbol{\omega}'\|_\infty \leq \kappa, \ \forall \boldsymbol{\omega}' \in \boldsymbol{\omega}^*(\boldsymbol{\mu}').$$

Proof. Due to Lemma 1 and Corollary 1, we have that, for all $\kappa = \kappa(\boldsymbol{\mu}) > 0$, there exists β_κ such that:

$$\|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \beta_\kappa \implies \Big| \max_{x \in \mathcal{X}_F(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x) - \max_{x \in \mathcal{X}_F(\boldsymbol{\mu}')} D(\boldsymbol{\mu}', \boldsymbol{\omega}, \neg x) \Big| \leq \kappa, \ \forall \boldsymbol{\omega} \in \Delta_K,$$
$$\|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \beta_\kappa \implies \inf_{\boldsymbol{\omega} \in \boldsymbol{\omega}^*(\boldsymbol{\mu})} \|\boldsymbol{\omega} - \boldsymbol{\omega}'\|_\infty \leq \kappa, \ \forall \boldsymbol{\omega}' \in \boldsymbol{\omega}^*(\boldsymbol{\mu}').$$

Recall that, as a consequence of Lemma 30, it holds that, for all $\epsilon > 0$, there exists $T_\epsilon : \forall t \geq T_\epsilon$, then, on $\mathcal{E}(t)$, $\|\boldsymbol{\mu} - \boldsymbol{\mu}'\|_\infty \leq \epsilon$ for all $s \geq h(t)$ and $\boldsymbol{\mu}' \in C_s$.

Picking $\epsilon = \beta_\kappa$ concluds the proof. □

The following lemma analyzes the behavior of the framework when the sequence of answers is not a converging one.

Lemma 40. *Let $\kappa > 0$ and let $T_\kappa \in \mathbb{N}$ as in Lemma 39. Then, for $t \geq T_\kappa$, under $\mathcal{E}(t)$, it holds that:*

$$\frac{1}{t} \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) \geq \frac{(t - \lceil h(t) \rceil)}{t} \min_{\boldsymbol{\omega} \in \mathrm{conv}(\boldsymbol{\omega}^*(\boldsymbol{\mu}))} \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{\omega}, \neg x) - \\ - \kappa(1 + C_{\boldsymbol{\mu}}) - \frac{K(1+\sqrt{t})}{t} C_{\boldsymbol{\mu}},$$

where $C_{\boldsymbol{\mu}} \geq 0$ is a problem dependent constant.

Proof. Consider $\frac{1}{t} \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x)$. By adding and subtracting $\frac{1}{t} \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{N}(t), \neg x)$ we have that

$$\frac{1}{t} \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) = \frac{1}{t} \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{N}(t), \neg x) - h_1(t)$$

where $h_1(t)$ is given by:

$$h_1(t) \coloneqq \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D\left(\boldsymbol{\mu}, \frac{\boldsymbol{N}(t)}{t}, \neg x\right) - \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D\left(\hat{\boldsymbol{\mu}}(t), \frac{\boldsymbol{N}(t)}{t}, \neg x\right).$$

Now, using Lemma 39 and noticing that $\boldsymbol{\mu} \in C_t$ on $\mathcal{E}(t)$, we have that $h_1(t) \leq \kappa$, thus leading to:

$$\frac{1}{t} \max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) \geq \frac{1}{t} \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{N}(t), \neg x) - \kappa.$$

We now continue by lower bounding $\frac{1}{t} \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{N}(t), \neg x)$. Specifically, for all $x \in \mathcal{X}^\star(\boldsymbol{\mu})$, we have that:

$$\begin{aligned}
&\frac{1}{t} \max_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} D(\boldsymbol{\mu}, \boldsymbol{N}(t), \neg x) \geq \frac{1}{t} D(\boldsymbol{\mu}, \boldsymbol{N}(t), \neg x)) \\
&= \frac{1}{t} \inf_{\boldsymbol{\lambda} \in \neg x} \sum_{k \in K} N_k(t) d(\mu_k, \lambda_k) \\
&= \frac{1}{t} \inf_{\boldsymbol{\lambda} \in \neg x} \left\{ \sum_{k \in K} \sum_{s \in [t]} \omega_k(s) d(\mu_k, \lambda_k) + \sum_{k \in K} \left(N_k(t) - \sum_{s \in [t]} \omega_k(s) \right) d(\mu_k, \lambda_k) \right\} \\
&\geq \frac{1}{t} \inf_{\boldsymbol{\lambda} \in \neg x} \sum_{k \in K} \sum_{s \in [t]} \omega_k(s) d(\mu_k, \lambda_k) + \frac{1}{t} \inf_{\boldsymbol{\lambda} \in \neg x} \sum_{k \in K} \left(N_k(t) - \sum_{s \in [t]} \omega_k(s) \right) d(\mu_k, \lambda_k) \\
&= \frac{1}{t} D\left(\boldsymbol{\mu}, \sum_{s=1}^{t} \boldsymbol{\omega}(s), \neg x\right) - h_2(t)
\end{aligned}$$

where $h_2(t)$ is given by:

$$\begin{aligned}
h_2(t) &\coloneqq \frac{1}{t} \inf_{\boldsymbol{\lambda} \in \neg x} \sum_{k \in [K]} \left(\sum_{s=1}^{t} \omega_k(s) - N_k(t) \right) d(\mu_k, \lambda_k) \\
&\leq \frac{K(1+\sqrt{t})}{t} \sup_{x \in \mathcal{X}^\star(\boldsymbol{\mu})} \inf_{\boldsymbol{\lambda} \in \neg x} \sum_{k \in [K]} d(\mu_k, \lambda_k) \\
&\leq \frac{K(1+\sqrt{t})}{t} C_{\boldsymbol{\mu}},
\end{aligned}$$

where in the last step we have used the fact that the exponential family is regular and bounded.

At this point, focus on $\frac{1}{t} D\left(\boldsymbol{\mu}, \sum_{s=1}^{t} \boldsymbol{\omega}(s), \neg x\right)$ and let us analyze $\boldsymbol{\omega}(s)$. We know that, $\boldsymbol{\omega}(s) \in \boldsymbol{\omega}^*(\boldsymbol{\mu}'_s)$ for some $\boldsymbol{\mu}'_s \in C_s$ (indeed, x_s is such that $x_s \in \mathcal{X}_F(\boldsymbol{\mu}'_s)$). Now, from Lemma 39, we know that, on $\mathcal{E}(t)$, there

exists $\{\bar{\boldsymbol{\omega}}(s)\}_{s\geq h(t)}$ such that (i) $\|\bar{\boldsymbol{\omega}}(s)-\boldsymbol{\omega}(s)\|_\infty \leq \kappa$ and (ii) $\bar{\boldsymbol{\omega}}(s)\in\boldsymbol{\omega}^*(\boldsymbol{\mu})$. This is direct by taking $\bar{\boldsymbol{\omega}}(s)\in \operatorname{argmin}_{\boldsymbol{\omega}\in\boldsymbol{\omega}^*(\boldsymbol{\mu})}\|\boldsymbol{\omega}(s)-\boldsymbol{\omega}\|_\infty$. Therefore, we obtain:

$$\begin{aligned}
\frac{1}{t}D\left(\boldsymbol{\mu},\sum_{s=1}^t\boldsymbol{\omega}(s),\neg x\right) &\geq \frac{1}{t}D\left(\boldsymbol{\mu},\sum_{s=\lceil h(t)\rceil}^t\boldsymbol{\omega}(s),\neg x\right)\\
&\geq \frac{t-\lceil h(t)\rceil}{t}D\left(\boldsymbol{\mu},\frac{\sum_{s=\lceil h(t)\rceil}^t\bar{\boldsymbol{\omega}}(s)}{t-\lceil h(t)\rceil},\neg x\right)+\frac{1}{t}D\left(\boldsymbol{\mu},\sum_{s=\lceil h(t)\rceil}^t\boldsymbol{\omega}(s)-\bar{\boldsymbol{\omega}}(s),\neg x\right)\\
&\geq \frac{t-\lceil h(t)\rceil}{t}D\left(\boldsymbol{\mu},\frac{\sum_{s=\lceil h(t)\rceil}^t\bar{\boldsymbol{\omega}}(s)}{t-\lceil h(t)\rceil},\neg x\right)-h_3(t)
\end{aligned}$$

where $h_3(t)$ is given by:

$$\begin{aligned}
h_3(t) &\coloneqq \frac{1}{t}\min_{x\in\mathcal{X}^\star(\boldsymbol{\mu})}\inf_{\boldsymbol{\lambda}\in\neg x}\sum_{s=\lceil h(t)\rceil}^t\sum_{k\in[K]}(\bar{\omega}_k(s)-\omega_k(s))d(\mu_k,\lambda_k)\\
&\leq \kappa\min_{x\in\mathcal{X}^\star(\boldsymbol{\mu})}\inf_{\boldsymbol{\lambda}\in\neg x}d(\mu_k,\lambda_k)\\
&\leq \kappa C_{\boldsymbol{\mu}},
\end{aligned}$$

Therefore, for all $x\in\mathcal{X}^\star(\boldsymbol{\mu})$ we have that:

$$\frac{1}{t}\max_{x\in\mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))}D(\hat{\boldsymbol{\mu}}(t),\boldsymbol{N}(t),\neg x)\geq\frac{t-\lceil h(t)\rceil}{t}D\left(\boldsymbol{\mu},\frac{\sum_{s=\lceil h(t)\rceil}^t\bar{\boldsymbol{\omega}}(s)}{t-\lceil h(t)\rceil},\neg x\right)-\kappa(1+C_{\boldsymbol{\mu}})-\frac{K(1+\sqrt{t})}{t}C_{\boldsymbol{\mu}}.$$

Hence:

$$\frac{1}{t}\max_{x\in\mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))}D(\hat{\boldsymbol{\mu}}(t),\boldsymbol{N}(t),\neg x)\geq\max_{x\in\mathcal{X}^\star(\boldsymbol{\mu})}\frac{t-\lceil h(t)\rceil}{t}D\left(\boldsymbol{\mu},\frac{\sum_{s=\lceil h(t)\rceil}^t\bar{\boldsymbol{\omega}}(s)}{t-\lceil h(t)\rceil},\neg x\right)-\kappa(1+C_{\boldsymbol{\mu}})-\frac{K(1+\sqrt{t})}{t}C_{\boldsymbol{\mu}}.$$

At this point, we note that $\frac{1}{t-\lceil h(t)\rceil}\sum_{s=\lceil h(t)\rceil}^t\bar{\boldsymbol{\omega}}(s)$ is an average of points that belongs to $\boldsymbol{\omega}^\star(\boldsymbol{\mu})$. Thus, this average belong to convex hull of $\boldsymbol{\omega}^\star(\boldsymbol{\mu})$. Consequently, we obtain:

$$\frac{1}{t}\max_{x\in\mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))}D(\hat{\boldsymbol{\mu}}(t),\boldsymbol{N}(t),\neg x)\geq\frac{t-\lceil h(t)\rceil}{t}\min_{\boldsymbol{\omega}\in\operatorname{conv}(\boldsymbol{\omega}^\star(\boldsymbol{\mu}))}\max_{x\in\mathcal{X}^\star(\boldsymbol{\mu})}D\left(\boldsymbol{\mu},\boldsymbol{\omega},\neg x\right)-\kappa(1+C_{\boldsymbol{\mu}})-\frac{K(1+\sqrt{t})}{t}C_{\boldsymbol{\mu}}.$$

thus concluding the proof. □

Finally, we are able to prove to prove the following result, Theorem 4, on the performance of the presented framework whenever $\{x_t\}$ is not a converging sequence.

Theorem 4. *Suppose that $\boldsymbol{\mu}\rightrightarrows\mathcal{X}_F(\boldsymbol{\mu})$ is not single-valued. Then, the presented framework δ-correct, and it always holds that:*

$$\limsup_{\delta\to0}\frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)}\leq\max_{\boldsymbol{\omega}\in\operatorname{conv}(\boldsymbol{\omega}^\star(\boldsymbol{\mu}))}\min_{x\in\mathcal{X}^\star(\boldsymbol{\mu})}\frac{1}{D(\boldsymbol{\mu},\boldsymbol{\omega},\neg x)}.$$

Proof. Let us define $\tilde{T}^*(\boldsymbol{\mu})$ as follows:

$$\tilde{T}^*(\boldsymbol{\mu})=\max_{\boldsymbol{\omega}\in\operatorname{conv}(\boldsymbol{\omega}^\star(\boldsymbol{\mu}))}\min_{x\in\mathcal{X}^\star(\boldsymbol{\mu})}\frac{1}{D(\boldsymbol{\mu},\boldsymbol{\omega},\neg x)}.$$

Let $\kappa>0$. From Lemma 40, we have that, for $t\geq T_\kappa$, on $\mathcal{E}(t)$:

$$\frac{1}{t}\max_{x\in\mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))}D(\hat{\boldsymbol{\mu}}(t),\boldsymbol{N}(t),\neg x)\geq\frac{(t-\lceil h(t)\rceil)}{t}\tilde{T}^*(\boldsymbol{\mu})^{-1}-\kappa(1+C_{\boldsymbol{\mu}})-\frac{K(1+\sqrt{t})}{t}C_{\boldsymbol{\mu}},\tag{26}$$

where $C_{\boldsymbol{\mu}}$ is a problem-dependent constant.

Let $\gamma \in \left(0, \frac{\tilde{T}^*(\boldsymbol{\mu})^{-1}}{4}\right]$, and take $\kappa = \kappa(\gamma) \leq \min\left(\frac{\gamma}{4}, \frac{\gamma}{4C_{\boldsymbol{\mu}}}\right)$. Furthermore, consider T_γ as the smallest $n \in \mathbb{N}$ such that $\frac{h(n)}{n}\tilde{T}^*(\boldsymbol{\mu})^{-1} \leq \frac{\gamma}{4}$ and $\frac{K(1+\sqrt{n})}{n}C_{\boldsymbol{\mu}} \leq \frac{\gamma}{4}$. Then, for $t \geq T_\gamma + T_{\kappa(\gamma)}$, Equation (26) implies that

$$\frac{1}{t}\max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) \geq \tilde{T}^*(\boldsymbol{\mu})^{-1} - \gamma.$$

Applying Lemma 42 with $\alpha = L = \gamma$ and $D = \tilde{T}^*(\boldsymbol{\mu})$, we have that, for $t \geq T_0(\gamma, \gamma, \delta, \tilde{T}^*(\boldsymbol{\mu})^{-1}) + T_\gamma + T_{\kappa(\gamma)}$:

$$\max_{x \in \mathcal{X}^\star(\hat{\boldsymbol{\mu}}(t))} D(\hat{\boldsymbol{\mu}}(t), \boldsymbol{N}(t), \neg x) \geq \beta_{t,\delta},$$

which implies that, on the good event, the algorithm stops at most at time $\lceil T_0(\gamma, \gamma, \delta, \tilde{T}^*(\boldsymbol{\mu})^{-1}) + T_\gamma + T_{\kappa(\gamma)} \rceil$.

Moreover, from Lemma 29, we obtain that:

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta] \leq T_0(\gamma, \gamma, \delta, \tilde{T}^*(\boldsymbol{\mu})^{-1}) + T_\gamma + T_{\kappa(\gamma)} + 1 + \sum_{s=0}^{+\infty} \mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}(s)^c).$$

From Lemma 19 in Garivier and Kaufmann [2016], $\sum_{s=0}^{+\infty} \mathbb{P}_{\boldsymbol{\mu}}(\mathcal{E}(s)^c)$ is finite. Thus, using the expression of $T_0(\gamma, \gamma, \delta, \tilde{T}^*(\boldsymbol{\mu})^{-1})$ given by Lemma 42, we have that:

$$\limsup_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]}{\log(1/\delta)} \leq \frac{1}{\tilde{T}^*(\boldsymbol{\mu})^{-1} - 2\gamma}.$$

Letting $\gamma \to 0$ concludes the proof. □

E.4.2 An example of the gap between Theorem 3 and Theorem 4

In the following, we provide a simple example where we can precisely evaluate the difference between Theorem 3 and Theorem 4.

Consider a problem with $K = 2$, $\boldsymbol{\mu} = (\mu_1, \mu_2)$ with Bernoulli distributions over $\mathcal{M} = [\alpha, 1-\alpha]$, and the task is to regress any of the two arms. Note that the task is non-trivial as it is intuitively more convenient to sample the arm with less variance. In this case, the set of correct answers is $\mathcal{X}^\star(\boldsymbol{\mu}) = \{(x, i) : |\mu_i - x| \leq \epsilon\}$. Consider $\alpha, \epsilon \in (0, 1/10)$ and $\boldsymbol{\mu}$ such that $\mu_1 = \alpha$ and $\mu_2 = 1 - \alpha$. By symmetry of the KL divergence at α and $1-\alpha$, we have that the set of optimal weights are trivially $(1, 0)$ for the answer $(\mu_1, 1)$ and $(0, 1)$ for the answer $(\mu_2, 2)$.

In this case, thus, we have that:

$$T^\star(\boldsymbol{\mu}) = \frac{1}{d(\alpha, \alpha+\epsilon))} = \frac{1}{d(1-\alpha, 1-\alpha-\epsilon)}.$$

On the other hand, $\tilde{T}^\star(\boldsymbol{\mu})$ actually evaluates to $2T^\star(\boldsymbol{\mu})$ (*i.e.*, as the worst weights over the convex hull are clearly $(0.5, 0.5)$).

F On simple solutions

This section is dedicated to show why some simple solution to the infinite answer problems fails, *i.e.*, they incur in sub-optimal sample complexity.

More precisely:

- In Appendix F.1, we discuss why the intuitive approach of discretizing the answer space and to apply Sticky Track-and-Stop on the resulting finite answer problem is not a good option to achieve statistical efficiency.
- In Appendix F.2, instead, we provide empirical evidence for the problem that Sticky Track-and-Stop faces when applied to an infinite answer space. This complements the discussion that we presented in Section 4, where we argued that the current proof of optimality of Sticky Track-and-Stop cannot be applied to the infinite answer setting. Indeed, the presence of infinite answers undermine the core argument behind optimality of Sticky Track-and-Stop (*i.e.*, the fact that it can stick to a correct answer in $\mathcal{X}_F(\boldsymbol{\mu})$).

F.1 Why discretizing the answer space does not work

In this section, we comment upon a simple approach that one might want to take when solving infinite answer problems on a compact, yet infinite, answer space. The simplest idea that one might have is to construct an α-net over the answer space and to apply Sticky Track-and-Stop to the resulting finite-answer problem. This solves the issue of sticking to a single-answer that we identified and discussed in Section 4. Nonetheless, as we now show, this automatically implies a loss in the statistical efficiency.

F.1.1 A simple example that explains the failure

We highlight the limits of this approach with a simple example. Consider the case where $K = 1$ and the target is the ϵ-regression of μ_1. In the following, we simply write μ since $K = 1$. Assume the distribution to be Gaussian with variance $1/2$ and let $\mathcal{M} = \mathcal{X} = [-1, 1]$. Then, it is easy to verify that Theorem 2 leads to $T^\star(\mu) = \frac{1}{\epsilon^2}$, and that this lower bound is only attained at $\mathcal{X}_F(\mu) = \mu$.

Now, denote by $\bar{\mathcal{X}}_\alpha$ the points in the α-net and suppose that $\alpha < \epsilon$. Then, for any $\mu \notin \bar{\mathcal{X}}_\alpha$, Sticky Track-and-Stop applied on the discretized answer space cannot select the statistically convenient answer $\mu \in \mathcal{X}_F(\mu)$ but can only select answers within $\bar{\mathcal{X}}_\alpha$. This implies that the algorithm can only attain an asymptotic upper bound of the form of $\frac{1}{(\epsilon-\alpha)^2}$, therefore failing to achieve optimality.

F.1.2 Abstracting away from the example

We observe that the previous scenario is only a very simple example and that discretizing "blindly" with an accuracy α can lead to more severe issues.

More generally, to apply an α-net and choose α so as to control the statistical error introduced by the discretization, one would need a Lipschitz continuity bound on the function $\frac{1}{D(\mu,\neg x)}$, and then take the minimum of such bound over all the possible $\boldsymbol{\mu} \in \mathcal{M}$. This is highly non-trivial: even in $D(\boldsymbol{\mu}, \neg x)$ (for which we can only prove a uniformly continuous result) the dependence in $\neg x$ enters through the (possibly non-convex) constraints of the underlying optimization problem.

Finally, any discretization must ensure that, for every $\boldsymbol{\mu} \in \mathcal{M}$, there exists some $x \in \bar{\mathcal{X}}_\alpha \cap \mathcal{X}^\star(\boldsymbol{\mu})$ such that $\boldsymbol{\mu} \notin \mathrm{cl}(\neg x)$. If this condition fails, Assumption 2 no longer holds in the resulting finite-answer problem, and the upper bounds for the discretized version would have infinite sample complexity.

F.2 Empirical Failure of Sticky Track-and-Stop

In the following, we present a simple example which highlights the core issue of Sticky Track-and-Stop in an infinite answer setting.

F.2.1 A simple example that explains the failure

Consider a bandit problem with $K = 4$ arms and Gaussian rewards with unitary variance. Let $\mathcal{M} = [-100, 100]^4$ and consider the problem of regressing up to an ϵ-accuracy (e.g., $\epsilon = 0.1$) either (μ_1, μ_2) or (μ_3, μ_4). Then, $\mathcal{X} = [-100, 100]^2$ and

$$\mathcal{X}^\star(\mu) = \{(x, y) \in \mathcal{X} : (|\mu_1 - x| \leq \epsilon \wedge |\mu_2 - y| \leq \epsilon) \vee (|\mu_3 - x| \leq \epsilon \wedge |\mu_4 - y| \leq \epsilon)\}.$$

Here, it is easy to verify that, for all $\boldsymbol{\lambda}$ such that $\mu_2 \ll \mu_4$, $\mathcal{X}_F(\boldsymbol{\lambda})$ directly evaluates to $\{(\lambda_1, \lambda_2)\} \cup \{(\lambda_3, \lambda_4)\}$ with corresponding oracle weights $(0.5, 0.5, 0, 0)$ and $(0, 0, 0.5, 0.5)$.

Consider the instance $\boldsymbol{\mu} = (0, -75, 0, 75)$ and the total order that first minimizes/maximizes the x-axis. Then, answers in $\mathcal{X}_t$ have approximately the following shape $[\hat{\mu}_1 \pm \mathrm{CI}_1, \hat{\mu}_2 \pm \mathrm{CI}_2]$ and $[\hat{\mu}_3 \pm \mathrm{CI}_3, \hat{\mu}_4 \pm \mathrm{CI}_4]$. Now, we note that:

- These two sets are well-separated by assumption on the y-axis but almost overlapping on the x-axis.
- Answers of the first set are related to weights that samples the first two arms, while answers in the second set leads to sampling arms 3 and 4.

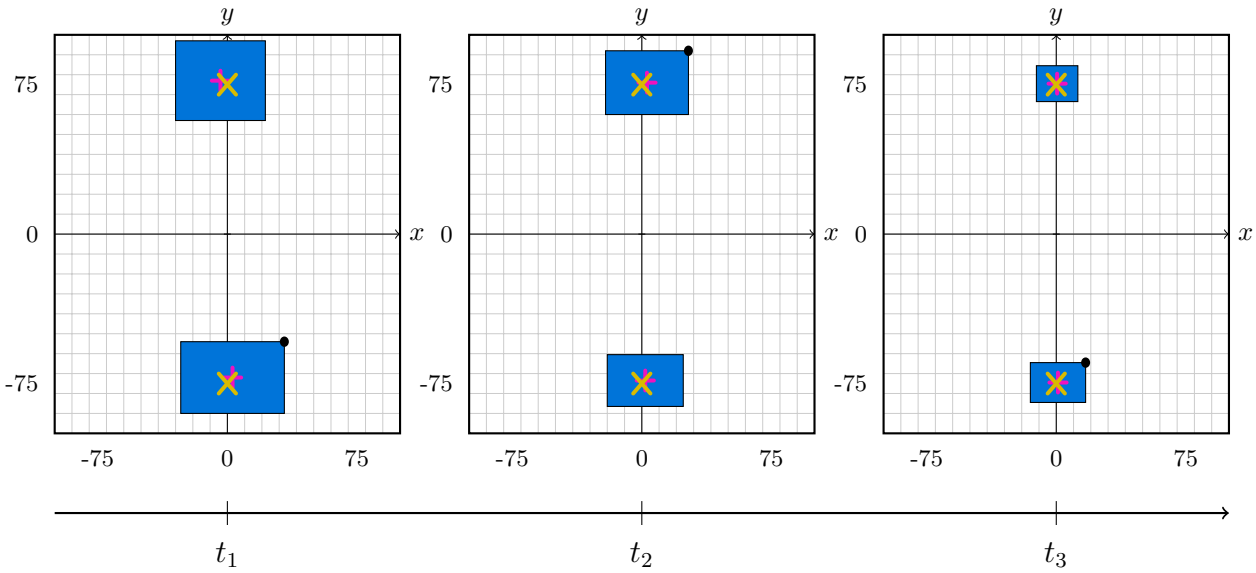

Figure 2: In **blue** we are visualizing how the set $\mathcal{X}_t$ is (approximately) behaving. The plus signs in **fuchsia** represents the empirical estimates of the first two arms and the last two arms that are generating $\mathcal{X}_t$. The **yellow** crosses represents the two unique answers in $\mathcal{X}_F(\boldsymbol{\mu})$. Finally, the **black** dot is the answer x_t selected by Sticky Track-and-Stop as it first maximizes the x-axis and then maximizes the y-axis. As a result, the algorithm oscillates between picking answers whose oracle weights are different.

As a consequence, since Sticky-TaS is using the total order that first minimizes over the x-axis, it will switch between picking answers in x_t from the two sets, and the corresponding expected pull proportions for small δ lies on the convex hull of $\boldsymbol{\omega}^\star(\boldsymbol{\mu})$. We have provided a visualization of this behavior in Figure 2.

F.2.2 Numerical Simulations

A part from the intuition that we provided above, we have also verified this issue of Sticky Track-and-Stop numerically. We have run the experiments on 8 CPUs Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz and 8 GB of RAM. In the simulation, we used $\delta = 10^{-80}$. We repeated the experiment 100 times and we compared the results with the ones of Sticky Sequence Track-and-Stop. Since $\mathcal{X}_F(\boldsymbol{\mu})$ is finite, we implemented the version that picks the next answer as $x_t \in \operatorname{argmin}_{x \in \mathcal{X}_t} \|x - x_{t-1}\|_\infty$.

The results are presented in Figure 3. As one can see from Figure 3 (*left*), the empirical sample complexity of Sticky Track-and-Stop is way higher than the one of Sticky Sequence Track-and-Stop, which, instead, is approaching the lower bound at $T^\star(\boldsymbol{\mu}) \log(1/\delta)$ (the red line in the plot). The issue, as we anticipated above, is that Sticky Track-and-Stop ends up in playing in the convex hull of $\boldsymbol{\omega}^\star(\boldsymbol{\mu})$. To verify this, in Figure 3 (*right*), we report the l_1 distance of the empirical pull proportions at the stopping time τ_δ from the closest $\boldsymbol{\omega} \in \boldsymbol{\omega}^\star(\boldsymbol{\mu})$; more formally $\inf_{\boldsymbol{\omega} \in \boldsymbol{\omega}^\star(\boldsymbol{\mu})} \|\boldsymbol{N}(\tau_\delta)/\tau_\delta - \boldsymbol{\omega}\|_1$. As one can see, Sticky Sequence Track-and-Stop ends up playing in expectation with proportions that are very close to $(0.5, 0.5, 0, 0)$ or $(0, 0, 0.5, 0.5)$. Sticky Track-and-Stop, on the other hand, will select answers in an oscillating fashion that leads to sub-optimal behavior.

F.2.3 Concluding Remarks

We conclude with two comments:

- The issue on the instance $\boldsymbol{\mu}$ that we considered might be solved by considering the total order that first minimizes/maximizes over y. However, one could create an instance where $\mu_2 = \mu_4$ while μ_1 and μ_3 are well-separated.
- We presented the problem in the simple setting of Gaussian distributions, for which an algorithm that samples arms 1 and 2 in equal proportion would suffice. Nonetheless, in the Bernoulli case no such fixed-proportion oracle exists and an adaptive algorithm becomes necessary (as sampling is more convenient where

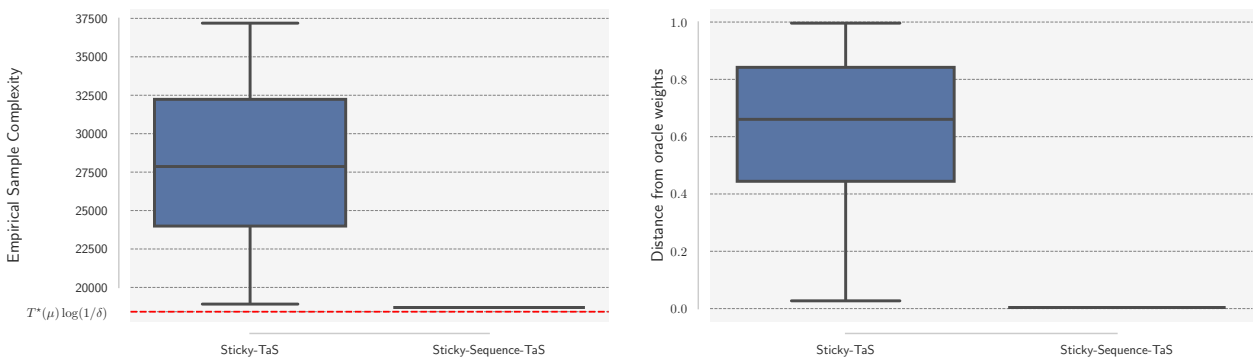

Figure 3: Empirical sample complexity of Sticky Track-and-Stop and Sticky Sequence Track-and-Stop on the presented regression task (*left*) and empirical l_1 distance of the empirical pull proportions at stopping from the closest $\boldsymbol{\omega} \in \boldsymbol{\omega}^\star(\boldsymbol{\mu})$; more formally $\inf_{\boldsymbol{\omega}\in\boldsymbol{\omega}^\star(\boldsymbol{\mu})} \|\boldsymbol{N}(\tau_\delta)/\tau_\delta - \boldsymbol{\omega}\|_1$ (*right*).

the variance is lower). Although the extension incurs additional mathematical complexity, our example can be generalized to cover this Bernoulli scenario. Specifically, we can pick μ_1 and μ_3 equal to α, for some small α, and $\mu_3 = 0.5 + \alpha$ and $\mu_4 = 0.5 - \alpha$. Let $\epsilon \ll \alpha$. Here, due to the symmetry of the KL divergence around 0.5, we have that regressing arm 1 and 2 is equivalent to regressing 3 and 4. Moreover, answers in $\mathcal{X}_t$ will still have the property of splitting into two distinct regions over the y-axis and the total order used by Sticky Track-and-Stop will lead to the same issue.

G Helper Lemmas

G.1 Tracking

The following lemma is a standard result related to the C-Tracking sampling rule. Tighter constant dependencies can be obtained using more fine-grained analysis Degenne et al. [2020b].

Lemma 41 (Lemma 7 in Garivier and Kaufmann [2016])**.** *For all $k \in [K]$ and all $t \geq 1$ it holds that $N_k(t) \geq \sqrt{t + K^2} - 2K$ and $\max_{k\in[K]} |N_k(t) - \sum_{s=1}^t \omega_k(s)| \leq K(1 + \sqrt{t})$.*

G.2 Stopping time

Similarly, the following lemma is useful in controlling the stopping time of the proposed algorithms. This result generalizes Lemma 1 in Ménard [2019].

Lemma 42. *Consider $D \geq 0$. Consider $\alpha > 0$ and $L \in \mathbb{R}$ such that $D - \alpha - L > 0$. There exists $C_\alpha > 0$ such that, for:*

$$T \geq \max\left\{C_\alpha, \frac{\log(1/\delta) + K\log(4\log(1/\delta) + 1)}{D - \alpha - L}\right\} := T_0(\alpha, L, \delta, D)$$

it holds that $D - L \geq \frac{\beta_{T,\delta}}{T}$.

Proof. Let C_α be such that, for $T \geq C_\alpha$ it holds that $6K\log(\log(T)+3)+K\tilde{C} \leq \alpha T$. Then, for $T \geq T_0(\alpha, L, \delta, D)$, we have that:

$$\begin{aligned}\frac{\beta_{T,\delta}}{T} &= \frac{\log\left(\frac{1}{\delta}\right) + K\log\left(4\log\left(\frac{1}{\delta}\right) + 1\right) + 6K\log(\log(T) + 3) + K\tilde{C}}{T} \\ &\leq \frac{\log\left(\frac{1}{\delta}\right) + K\log\left(4\log\left(\frac{1}{\delta}\right) + 1\right)}{T} + \alpha \\ &\leq D - L.\end{aligned}$$

where (i) in the first step we have used the definition of $\beta_{t,\delta}$, i.e., Equation (15), (ii) in the second one the definition of C_α, and (iii) in the third one the definition of $T_0(\alpha, L, \delta, D)$. □

H Mathematical Background

This section contains mathematical background that can be helpful throughout the document. Specifically, Appendix H.1 shows that is possible to have an exact covering of a compact set. Appendix H.2 contains useful information on canonical exponential families. Appendix H.3 provides simple results on the infimum of optimization problems. Finally, Appendix H.4 presentes auxiliary results on correspondences and set-valued analysis.

H.1 Set Theory

Lemma 43 (Exact covering of a compact set)**.** *Let $\mathcal{X} \subset \mathbb{R}^d$ be a compact set. For all $\rho > 0$, there exists $n_\rho \in \mathbb{N}$ finite and $\{\mathcal{X}_i\}_{i=1}^{n_\rho}$ such that $\mathcal{X}_i \subset \mathbb{R}^d$ is compact, $\mathcal{X} = \bigcup_{i=1}^{n_\rho} \mathcal{X}_i$, and for all $i \in [n_\rho]$, $\mathcal{X}_i \subseteq \mathcal{B}_\rho(x_i)$ for some $x_i \in \mathcal{X}$.*

Proof. Let $\bar{\mathcal{B}}_\rho(x) = \{y \in \mathbb{R}^d : ||x - y||_\infty < \rho\}$ be an open ball of radius ρ centered in ρ. Then, for all $\rho > 0$, it holds that $\mathcal{X} \subset \bigcup_{x \in \mathcal{X}} \bar{\mathcal{B}}_\rho(x_i)$. Furthermore, since $\mathcal{X}$ is compact, every open cover of a compact set admits a finite subcover, that is, there exists $n_\rho \in \mathbb{N}$ finite and a collection of points $\{x_i\}_{i=1}^{n_\rho}$ such that $\mathcal{X} \subset \bigcup_{i=1}^{n_\rho} \bar{\mathcal{B}}_\rho(x_i)$, where each $x_i \in \mathcal{X}$ by construction. Thus, by taking the closure of each ball, we obtain, $\mathcal{X} \subset \bigcup_{i=1}^{n_\rho} \mathcal{B}_\rho(x_i)$. Finally, we have that:

$$\mathcal{X} = \bigcup_{i=1}^{n_\rho} (\mathcal{B}_\rho(x_i) \cap \mathcal{X}) \coloneqq \bigcup_{i=1}^{n_\rho} \mathcal{X}_i.$$

To conclude the proof, we note that the intersections of compact euclidean subsets is compact. Hence, $\mathcal{X}_i = \mathcal{B}_\rho(x_i) \cap \mathcal{X}$ is compact. □

H.2 Canonical Exponential Family

In a canonical and one-parameter exponential family, distributions are indexed according to a parameter $\phi \in \Phi$, and each distribution ν_ϕ is absolutely continuous with respect to a reference measure ρ on $\mathbb{R}$ such that:

$$\frac{\mathrm{d}\nu_\eta}{\mathrm{d}\rho}(x) = \exp(x\eta - b(\eta)),$$

where $b : \Phi \to \mathbb{R}$ is a twice differentiable convex function. Each distribution ν_ϕ can be uniquely identified with its mean μ, which is given by $\dot{b}(\eta)$. Given an interval of open means (i.e., the family is regular), b is strictly convex on that interval, and the distribution is non-degenerate, meaning that its variance is strictly positive. The KL distribution between two distributions $\nu_\eta, \nu_{\eta'}$ with means μ, μ', is given by:

$$\mathrm{KL}(\nu_\eta, \nu_{\eta'}) = d(\mu, \mu') = b(\eta') - b(\eta) - \dot{b}(\eta)(\eta' - \eta).$$

Now, consider two bandits $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$. After t rounds it holds that

$$\begin{aligned}
\ln \frac{\mathrm{d}\mathbb{P}_{\boldsymbol{\mu}}}{\mathrm{d}\mathbb{P}_{\boldsymbol{\lambda}}} &= \ln \frac{\mathrm{d}\mathbb{P}_{\hat{\boldsymbol{\mu}}(t)}}{\mathrm{d}\mathbb{P}_{\boldsymbol{\lambda}}} - \ln \frac{\mathrm{d}\mathbb{P}_{\hat{\boldsymbol{\mu}}(t)}}{\mathrm{d}\mathbb{P}_{\boldsymbol{\mu}}} \\
&= \sum_{k \in [K]} N_k(t) \left(d(\hat{\mu}_k(t), \lambda_k) - d(\hat{\mu}_k(t), \mu_k) \right) \\
&= \sum_{k \in [K]} N_k(t) \Big(d(\mu_k, \lambda_k) + (\tilde{\lambda}_k - \tilde{\mu}_k)(\mu_k - \hat{\mu}_k(t)) \Big),
\end{aligned}$$

where $\tilde{\mu}_k = \dot{b}^{-1}(\mu_k)$ and $\tilde{\lambda}_k = \dot{b}^{-1}(\lambda_k)$ represents the natural parameter of the distributions with mean λ_k and μ_k respectively. We also recall that $\sum_{k \in [K]} N_k(t)(\tilde{\lambda}_k - \tilde{\mu}_k)(\mu_k - \hat{\mu}_k(t)))$ is a martingale. We also recall that in the third equality above, we used the following property of canonical exponential family.

Lemma 44 (KL difference in canonical exponential families)**.** *Consider three distributions in a one-dimensional exponential family with means a, b, c. Then, it holds that:*

$$d(a, b) = d(a, c) + d(c, b) + (\tilde{b} - \tilde{c})(c - a).$$

Proof. For a proof, see e.g., Lemma E.6 in Poiani et al. [2025]. □

Finally, we show lipschitzianity properties of the KL divergence and natural parameters when dealing with canonical exponential family. This result is well-known and we report a proof for completeness.

Lemma 45 (Local lipschitzianity of KL divergence and natural parameters)**.** *Consider $\mu, \lambda \in \Theta$, and denote by $\tilde{\mu}, \tilde{\lambda}$ their parameter in the canonical exponential family. Then, it holds that:*

$$|\tilde{\lambda} - \tilde{\mu}| \leq C_{1,\mu,\lambda}|\lambda - \mu|$$
$$d(\mu, \lambda) \leq C_{2,\mu,\lambda}(\lambda - \mu)^2,$$

where,

$$C_{1,\mu,\lambda} := \frac{1}{\min_{\xi \in [\min\{\tilde{\mu},\tilde{\lambda}\},\max\{\tilde{\mu},\tilde{\lambda}\}]} \ddot{b}(\xi)} \quad \text{and} \quad C_{2,\mu,\lambda} := \frac{C_{1,\mu,\lambda}}{2} \max_{\xi \in [\min\{\tilde{\mu},\tilde{\lambda}\},\max\{\tilde{\mu},\tilde{\lambda}\}]} \ddot{b}(\xi).$$

Proof. We first recall that $b(\cdot)$ is convex and twice differentiable. Therefore, we have that:

$$b(\tilde{\lambda}) \leq b(\tilde{\mu}) + \mu(\tilde{\lambda} - \tilde{\mu}) + (\tilde{\lambda} - \tilde{\mu})^2 \max_{\xi \in [\min\{\tilde{\mu},\tilde{\lambda}\},\max\{\tilde{\mu},\tilde{\lambda}\}]} \frac{\ddot{b}(\xi)}{2}.$$

Plugging this result within $d(\mu, \lambda)$, we obtain:

$$d(\mu, \lambda) \leq (\tilde{\lambda} - \tilde{\mu})^2 \max_{\xi \in [\min\{\tilde{\mu},\tilde{\lambda}\},\max\{\tilde{\mu},\tilde{\lambda}\}]} \frac{\ddot{b}(\xi)}{2}. \tag{27}$$

We now recall that $\dot{b}(\tilde{\mu}) = \mu$. Suppose that $\tilde{\mu} < \tilde{\lambda}$. Then, by the mean value theorem of integration, we have that:

$$\ddot{b}(\tilde{c}) = \frac{\lambda - \mu}{\tilde{\lambda} - \tilde{\mu}} \text{ for } \tilde{c} \in [\tilde{\mu}, \tilde{\lambda}]. \implies \tilde{\lambda} - \tilde{\mu} = \frac{\lambda - \mu}{\ddot{b}(\tilde{c})}.$$

Similarly, when $\tilde{\lambda} < \tilde{\mu}$, we have:

$$\ddot{b}(\tilde{c}) = \frac{\mu - \lambda}{\tilde{\mu} - \tilde{\lambda}} \text{ for } \tilde{c} \in [\tilde{\lambda}, \tilde{\mu}]. \implies \tilde{\mu} - \tilde{\lambda} = \frac{\mu - \lambda}{\ddot{b}(\tilde{c})}.$$

Chaining these results, we obtain:

$$(\tilde{\lambda} - \tilde{\mu})^2 \leq \frac{(\lambda - \mu)^2}{\left(\min_{\xi \in [\min\{\tilde{\mu},\tilde{\lambda}\},\max\{\tilde{\mu},\tilde{\lambda}\}]} \ddot{b}(b^{-1}(\xi))\right)^2}.$$

Notice that, since Θ is an open interval, that min is well-defined and different from zero.

Plugging this upper bound within Equation 27, we can conclude the proof:

$$d(\mu, \lambda) \leq (\lambda - \mu)^2 \frac{\max_{\xi \in [\min\{\tilde{\mu},\tilde{\lambda}\},\max\{\tilde{\mu},\tilde{\lambda}\}]} \ddot{b}(\xi)}{2\left(\min_{\xi \in [\min\{\tilde{\mu},\tilde{\lambda}\},\max\{\tilde{\mu},\tilde{\lambda}\}]} \ddot{b}(\xi)\right)^2}.$$

□

As a corollary of Lemma 45, we have that the parameter space and the KL divergence are Lipschitz on any compact subset of the parameter space Θ. Again, this result is well-known and we report a proof for completeness. We remark that, since we assumed Θ to be strictly contained in an open interval, then, we enjoy this stronger Lipschitzianity result.

Corollary 2 (Lipschitzianity over a compact set)**.** *Let $\widetilde{\Theta} \subset \Theta$ be a compact set. Then, there exists constants $C_1, C_2 > 0$ such that, for all $\mu, \lambda \in \widetilde{\Theta}$, it holds that:*

$$|\tilde{\mu} - \tilde{\lambda}| \leq C_1 |\mu - \lambda| \quad \text{and} \quad d(\mu, \lambda) \leq C_2 (\lambda - \mu)^2.$$

Furthermore, there exists $D_1, D_2 > 0$ such that $|\tilde{\mu} - \tilde{\lambda}| < D_1$ and $d(\mu, \lambda) < D_2$.

Proof. Since $\widetilde{\Theta}$ is compact, there exists $\mu_{\min}, \mu_{\min}$ s.t. $\mu \in [\mu_{\min}, \mu_{\max}]$ for all $\mu \in \widetilde{\Theta}$. In particular, let $\mu_{\min} = \min_{\mu \in \widetilde{\Theta}} \mu$ and $\mu_{\max} = \max_{\mu \in \widetilde{\Theta}} \mu$. Denote by $\widetilde{\Phi} = \{\dot{b}^{-1}(\mu) : \forall \mu \in [\mu_{\min}, \mu_{\max}]\}$. Since $[\mu_{\min}, \mu_{\max}]$ is compact and $\dot{b}^{-1}(\cdot)$ is a continuous function, $\widetilde{\Phi}$ is compact as well.

Now, from Lemma 44, we know that, for all $\mu, \lambda \in \widetilde{\Theta}$, it holds that:

$$\begin{aligned}
|\tilde{\lambda} - \tilde{\mu}| &\leq \frac{|\lambda - \mu|}{\min_{\xi \in [\min\{\tilde{\mu}, \tilde{\lambda}\}, \max\{\tilde{\mu}, \tilde{\lambda}\}]} \ddot{b}(\xi)} \\
&\leq \frac{|\lambda - \mu|}{\inf_{\xi \in \widetilde{\Phi}} \ddot{b}(\xi)} \\
&= \frac{|\lambda - \mu|}{\min_{\xi \in \widetilde{\Phi}} \ddot{b}(\xi)} \\
&\coloneqq C_1 |\lambda - \mu|
\end{aligned}$$

where the inf can be replaced with a min since the optimization set is compact and $\ddot{b}(\cdot)$ is continuous over $\widetilde{\Theta}$. This is due to the fact that $\widetilde{\Theta} \subset \Theta$ and Θ is an open interval. Hence, the exponential family is regular and the function $b(\cdot)$ is $\mathcal{C}^\infty$, see Theorem 5.8 in Lehmann and Casella [2006]. Notice, furthermore, that the regularity of the exponential family also implies that $\ddot{b}$ is strictly positive over the considered domain. Finally, taking $D_1 \coloneqq C_1 |\mu_{\max} - \mu_{\min}|$ shows that $|\tilde{\mu} - \tilde{\lambda}|$ is bounded.

We now analyze divergence $d(\cdot, \cdot)$ using similar arguments. From Lemma 44, we have that:

$$\begin{aligned}
d(\mu, \lambda) &\leq \frac{\max_{\xi \in [\min\{\tilde{\mu}, \tilde{\lambda}\}, \max\{\tilde{\mu}, \tilde{\lambda}\}]} \ddot{b}(\xi)}{\min_{\xi \in [\min\{\tilde{\mu}, \tilde{\lambda}\}, \max\{\tilde{\mu}, \tilde{\lambda}\}]} \ddot{b}(\xi)} (\lambda - \mu)^2 \\
&\leq \frac{\sup_{\xi \in \widetilde{\Theta}} \ddot{b}(\xi)}{\inf_{\xi \in \widetilde{\Theta}} \ddot{b}(\xi)} (\lambda - \mu)^2 \\
&= \frac{\max_{\xi \in \widetilde{\Theta}} \ddot{b}(\xi)}{\min_{\xi \in \widetilde{\Theta}} \ddot{b}(\xi)} (\lambda - \mu)^2 \\
&\leq \frac{\max_{\xi \in \widetilde{\Theta}} \ddot{b}(\xi)}{\min_{\xi \in \widetilde{\Theta}} \ddot{b}(\xi)} (\mu_{\max} - \mu_{\min}) |\lambda - \mu| \\
&\coloneqq C_2 |\lambda - \mu|.
\end{aligned}$$

Finally, taking $D_2 \coloneqq C_2 |\mu_{\max} - \mu_{\min}|$ shows that $d(\mu, \lambda)$ is bounded. □

H.3 Results on the infimum

Here, we simply state that we can split the infimum by considering unions of the optimization sets.

Lemma 46. *Consider $\mathbb{X} \subseteq \mathbb{R}^d$, and $f : \mathbb{X} \to \mathbb{R}$. Let $\mathbb{X}_1, \mathbb{X}_2 : \mathbb{X}_1 \cup \mathbb{X}_2 = \mathbb{X}$. Then, it holds that:*

$$\inf_{x \in \mathbb{X}} f(x) = \min \left\{ \inf_{x \in \mathbb{X}_1} f(x), \inf_{x \in \mathbb{X}_2} f(x) \right\}.$$

H.4 Correspondences

We start with some preliminary definitions.

Let $\mathbb{U}$ be a topological space and let $U \in \mathbb{U}$ s.t. $U \neq \emptyset$. A function $f : U \to \mathbb{R}$ is inf-compact on U if the level sets

$$\mathcal{D}_f(\lambda, U) = \{y \in U : f(y) \leq \lambda\}$$

are compact for all $\lambda \in \mathbb{R}$. Furthermore, it is upper semicontinuous if all the strict level sets

$$\mathcal{D}_f^{<}(\lambda, U) = \{y \in U : f(y) < \lambda\}$$

are open.

Next, consider a correspondence $\phi : \mathbb{X} \rightrightarrows \mathbb{Y}$. Then, for any $Z \subseteq \mathbb{X}$, let

$$\mathrm{Gr}_Z(\phi) = \{(x, y) \in Z \times \mathbb{Y} : y \in \phi(x)\}.$$

Consider a function $f : \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$. Then, f is $\mathbb{K}$-inf-compact on $\mathrm{Gr}_{\mathbb{X}}(\phi)$, if all for all compact subsets K of $\mathbb{X}$, it holds that f is inf-compact on $\mathrm{Gr}_K(\phi)$.

We now state Berge's Maximum Theorem [Berge, 1959].

Theorem 5 (Berge's Maximum Theorem)**.** *Let $\mathbb{X}, \mathbb{Y}$ be Haussdorf topological spaces. Let $f : \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$ be a continuous function, and let $\phi : \mathbb{X} \rightrightarrows \mathbb{Y}$ be a continuous and compact-valued correspondence. Then, let $f^* : \mathbb{X} \to \mathbb{R}$ and $\phi^* : \mathbb{X} \rightrightarrows \mathbb{Y}$ be defined as follows:*

$$f^*(x) = \max_{y \in \phi(x)} f(x, y)$$
$$\phi^*(x) = \operatorname*{argmax}_{y \in \phi(x)} f(x, y).$$

Then, f^* is continuous over $\mathbb{X}$ and ϕ^* is upper hemicontinuous and compact-valued over $\mathbb{X}$.

Next, we introduce the following result that extends Berge's maximum theorem to non-compact and only lower hemicontinuous correspondences.

Theorem 6 ([Feinberg et al., 2014, Theorem 1.2])**.** *Let $\mathbb{X}$ be a compactly generated topological space and $\mathbb{Y}$ be a Hausdorff topological space. Let $\phi : \mathbb{X} \rightrightarrows \mathbb{Y}$ be a lower hemicontinuous correspondence, and let $f : \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$ be $\mathbb{K}$-inf-compact and upper semi-continuous on $\mathrm{Gr}_{\mathbb{X}}(\phi)$. Then, let $f^* : \mathbb{X} \to \mathbb{R}$ and $\phi^* : \mathbb{X} \rightrightarrows \mathbb{Y}$ be defined as follows:*

$$f^*(x) = \sup_{y \in \phi(x)} f(x, y)$$
$$\phi^*(x) = \operatorname*{argmax}_{y \in \phi(x)} f(x, y).$$

Then, f^* is continuous and and ϕ^* is upper hemicontinuous and compact-valued.

We recall that, when dealing with topological subspaces $\mathbb{X}$ of $\mathbb{R}^d$ with the inherited euclidean topology, then $\mathbb{X}$ are metric topological spaces, and hence, Haussdorf and compactly generated. Therefore, Theorem 5 and Theorem 6 can be applied.

Finally, we report the following results that we use to prove the continuity and compactness of (some) correspondence within our analysis.

Theorem 7 ([Aubin, 1999, Proposition 1.4.14])**.** *Let $\mathbb{X} \subseteq \mathbb{R}^{n_x}$, $\mathbb{Y} \subseteq \mathbb{R}^{n_y}$ and $\mathbb{Z} \subseteq \mathbb{R}^{n_z}$ be three sets and let $U : \mathbb{X} \rightrightarrows \mathbb{Z}$ be a compact-valued continuous correspondence. Then let $g : \mathrm{Graph}(U) \to \mathbb{Y}$ be a continuous function. Then the correspondence $G : \mathbb{X} \rightrightarrows \mathbb{Y}$ defined as $G(x) = \cup_{u \in U(x)}\{g(x, u)\}$ is continuous and compact valued.*

Theorem 8 ([Aubin, 1999, Proposition 1.4.8])**.** *The graph of an upper hemicontinuous set-valued map $F : \mathbb{X} \rightrightarrows \mathbb{Y}$ with closed domain and closed values is closed. The converse is true if we assume that $\mathbb{Y}$ is compact.*