

Geometry Meets Incentives: Sample-Efficient Incentivized Exploration with Linear Contexts

Benjamin Schiffer*

Mark Sellke†

June 3, 2025

Abstract

In the incentivized exploration model, a principal aims to explore and learn over time by interacting with a sequence of self-interested agents. It has been recently understood that the main challenge in designing incentive-compatible algorithms for this problem is to gather a moderate amount of initial data, after which one can obtain near-optimal regret via posterior sampling. With high-dimensional contexts, however, this *initial exploration* phase requires exponential sample complexity in some cases, which prevents efficient learning unless initial data can be acquired exogenously. We show that these barriers to exploration disappear under mild geometric conditions on the set of available actions, in which case incentive-compatibility does not preclude regret-optimality. Namely, we consider the linear bandit model with actions in the Euclidean unit ball, and give an incentive-compatible exploration algorithm with sample complexity that scales polynomially with the dimension and other parameters.

*Department of Statistics, Harvard University — *E-mail:* bschiffer1@g.harvard.edu.

†Department of Statistics, Harvard University — *E-mail:* msellke@fas.harvard.edu.

1 Introduction

The exploration/exploitation trade-off is fundamental to online decision making. This trade-off is classically exemplified by the multi-armed bandit problem, where a single agent chooses actions sequentially and learns to improve over time. In this setting, the agent has clear justification for early exploration because they can reap future rewards by exploiting the knowledge gained by exploration. But what if agents are unable to reap these future rewards, and so each decision must be justifiable on its own terms without taking into account future rewards? This may occur when actions are *recommendations* made to different agents by a central platform based on user feedback, as in e-commerce, traffic routing, movies, restaurants, etc. In these settings, the agents may have a prior for the rewards of the actions in addition to the recommendation made by the platform. Therefore, while the platform makes recommendations with the goal of learning over time, individual myopic agents will decline to follow any recommendations which seem suboptimal based on their individual priors.

The *incentivized exploration* problem was introduced in [KMP14, CH18] to understand this fundamental tension, and extends the well-studied problem of *Bayesian Persuasion* in information design [BM19, Kam19]. The model adopted in these early works consists of a finite set of actions, each with a (publicly shared) Bayesian prior distribution for rewards. A sequence of agents arrives one by one, and each agent is recommended an action by a central planner. The central planner’s recommendations are made using a (publicly known) randomized algorithm, and the agents are assumed to be selfish and rational, aiming only to maximize their own expected reward. Given the planner’s recommendation, each agent computes and chooses the posterior-optimal action (using Bayes’ rule). As observed in the original works, thanks to the revelation principle of [Mye86], the latter step is equivalent to assuming that the planner’s recommendations are *Bayesian incentive-compatible* (BIC) so that rational agents will always follow the recommendations (at least under the assumption of agent homogeneity).

Initially, most work on incentivized exploration dealt with small finite sets of actions [MSS20, MSSW22], exploring economic aspects of the problem such as exogenous payments for exploration [FKKK14, KKM⁺17, WH18, AT20, WXL⁺23], partial data disclosure [IMSW20], and agent heterogeneity [IMSW19]. See also the surveys [Sli19, Chapter 11] and [IEV23, Chapter 31]. More recently, extensions to combinatorial action sets, multi-stage reinforcement learning, and linear contexts were also considered in [HNSW22, SS24, Sel23]. In these more complex machine learning settings, a fundamental question is how the regret scales with problem parameters such as the size of the action space.

This quantitative dependence was studied in [SS23], which provided a new two-stage algorithm. The algorithm first obtains a constant number of samples from each arm, and then switches permanently to Thompson sampling. The initial stage enjoys sample-efficiency thanks to a carefully tuned “exponential exploration” strategy, and Thompson sampling is BIC after this constant amount of initial exploration (see [GKS24] for a more general perspective on the latter property). This yielded polynomial regret dependence on the number of actions and other natural parameters, and ensured the “price of incentivization” in the re-

gret is only additive relative to Thompson sampling, which is known to exhibit near-optimal performance [KKM12, BL13, AG17, RV14, RV16, ZL19, BS22]. Importantly, however, all of these results require that the rewards for the actions are independent under the prior.

A natural setting to consider dependent actions is the linear bandit model, where Thompson sampling remains a gold-standard algorithm [AG13, DVR18]. For this setting, [Sel23] showed that Thompson sampling is again BIC after an initial data collection stage under mild conditions, but that the sample complexity of *collecting* initial data can scale exponentially with the dimension. In some applications, using exogenous payments for initial exploration can bypass this exponential barrier, but such workarounds are contingent on problem-specific regulatory and ethical constraints. In our work, we show that the exponential barrier disappears when the action set is the d -dimensional unit ball, and we provide an incentive-compatible initial exploration algorithmic with polynomial sample complexity.

1.1 Our Results

We consider a linear bandit problem where the set \mathcal{A} of possible actions is the d -dimensional unit ball. At each time step $t \geq 1$, the algorithm chooses an action $\mathbf{A}^{(t)} \in \mathcal{A}$ and observes a noisy reward

$$r^{(t)} = \langle \mathbf{A}^{(t)}, \boldsymbol{\ell}^* \rangle + w_t,$$

where $w_t \sim N(0, 1)$ and $w_{t_1} \perp w_{t_2}$ for all $t_1 \neq t_2$. We assume that $\boldsymbol{\ell}^*$ is drawn from a known prior distribution μ on \mathbb{R}^d . As our model and algorithm will be rotationally invariant, we assume for convenience (without loss of generality) that $\mathbb{E}[\boldsymbol{\ell}_i^*] = 0$ for all $i > 1$ and $\mathbb{E}[\boldsymbol{\ell}_1^*] \geq 0$. Importantly, unlike in [SS23], we do *not* require that $\boldsymbol{\ell}_i^*$ is independent of $\boldsymbol{\ell}_j^*$ for $i \neq j$.

We are interested in exploration via Bayesian Incentive Compatible (BIC) algorithms, defined formally as follows. As discussed above, an algorithm being BIC implies that agents will follow the recommendations made by a platform using that algorithm.

Definition 1.1. An action $\mathbf{A}^{(t)}$ at time $t \geq 1$ is Bayesian Incentive Compatible (BIC) if for all $A \in \mathcal{A}$:

$$\mathbb{E}[\langle \boldsymbol{\ell}^*, A \rangle \mid \mathbf{A}^{(t)} = A] = \sup_{A' \in \mathcal{A}} \mathbb{E}[\langle \boldsymbol{\ell}^*, A' \rangle \mid \mathbf{A}^{(t)} = A].$$

Similarly, a bandit algorithm is BIC if for all $t \in [1, T]$ the algorithm recommends a BIC action $\mathbf{A}^{(t)}$.

Our main goal is to develop BIC algorithms that explore the entire action space using $\text{poly}(d)$ samples. Formally, we want our algorithm to choose actions $\mathbf{A}^{(1)}, \dots, \mathbf{A}^{(T)}$ such that for some constant $\lambda > 0$

$$\sum_{t=1}^T (\mathbf{A}^{(t)})^{\otimes 2} \succeq \lambda \mathbf{I}, \tag{1}$$

i.e. all eigenvalues of $\sum_{t=1}^T (\mathbf{A}^{(t)})^{\otimes 2}$ are at least λ . This was termed λ -spectral exploration in [Sel23], where it was shown to imply Thompson Sampling is BIC after time T under mild conditions.

For some priors, it is impossible to explore the action space with a BIC algorithm (see e.g. [MSS20, Section 4] or [SS23, Section 8]). Thus we will require mild non-degeneracy assumptions on the prior. Roughly speaking, ℓ^* should not be confined to any half-space and should have neither minuscule nor enormous fluctuations in any direction.

Assumption 1.2. *There exist constants $c_d, \epsilon_d, c_v, K > 0$ such that:*

1. ℓ^* is not confined to any half-space: $\min_{\|\mathbf{v}\|=1} \Pr(\langle \mathbf{v}, \ell^* \rangle \geq c_d) \geq \epsilon_d$.
2. ℓ^* has non-degenerate covariance: $\min_{\|\mathbf{v}\|=1} \text{Var}(\langle \mathbf{v}, \ell^* \rangle) \geq c_v$.
3. ℓ^* is sub-gaussian: $\max_{\|\mathbf{v}\|=1} \mathbb{P}(|\langle \mathbf{v}, \ell^* \rangle| \geq t) \leq 2e^{-t^2/K^2}$ for all $t > 0$.

The first two conditions above are both necessary in some form; see Appendix B. The third is a standard condition on the fluctuations of μ . Our main result is as follows.

Theorem 1.3. *Under Assumption 1.2, there exists a BIC algorithm (Algorithm 1) which almost surely achieves $\bar{\lambda}$ -spectral exploration in sample complexity*

$$\bar{\lambda} \left(\frac{d}{c_v + c_d} \right)^{O(1)} \log(1/\epsilon_d). \quad (2)$$

Note that (2) depends polynomially on c_d, c_v but only logarithmically on ϵ_d . This is important because in typical high-dimensional settings, ϵ_d will be exponentially small in the dimension while the other parameters will not. The next two propositions give illustrative but not exhaustive examples of such high-dimensional distributions. In the first, we take μ to be uniform on a convex body \mathcal{K} with $B_r(0) \subseteq \mathcal{K} \subseteq B_1(0)$ for some $0 < r < 1$, and say such μ is r -regular. This is the main setting considered in [Sel23], and as explained in Section 4, combining [Sel23] with our results yields an end-to-end low-regret algorithm which is ϵ -BIC (i.e. with ϵ subtracted from the right-hand side in Definition 1.1) with initial sample complexity $\text{poly}(d, 1/r, 1/\epsilon)$. (The combination only satisfies ϵ -BIC because [Sel23] only shows Thompson sampling is ϵ -BIC unless actions are well-separated.)

Proposition 1.4 (Proof in Appendix L). *Any r -regular μ satisfies Assumption 1.2 with $c_d = r/3$ and $\epsilon_d = (r/3)^d$ and $c_v = \Omega(r^2/d^2)$ and $K = 1.25$.*

The second example consists of log-concave distributions (e.g. Gaussians). We say a density $d\mu(x) \propto e^{-V(x)}dx$ is α -log-concave and β -log-smooth for $0 < \alpha \leq \beta$ if

$$-\beta I_d \preceq \nabla^2 V(\mathbf{x}) \preceq -\alpha I_d, \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

Proposition 1.5 (Proof in Appendix M). *Let μ be αd -log-concave and βd -log-smooth with mode \mathbf{x}^* satisfying $\|\mathbf{x}^*\| \leq \gamma$. Then Assumption 1.2 holds with $\epsilon_d \geq \frac{J e^{-J^2 \beta d/2} \sqrt{\alpha d}}{(1+J^2 \beta d) \sqrt{2\pi}}$ for $J = \gamma + \frac{2}{\sqrt{\alpha d}} + c_d$ and any $c_d > 0$. Further, we may take $c_v = \frac{1}{\beta d}$ and $K = 2(\gamma + \sqrt{\alpha d} + (\alpha d)^{-1/2})$. Conversely, there exists such μ with $\|\mathbf{x}^*\| = 0$ and $(\alpha, \beta) = (1, 2)$ in which $\min_{\|\mathbf{v}\|=1} \Pr(\langle \mathbf{v}, \ell^* \rangle \geq 0) \leq e^{-\Omega(d)}$.*

We note that if μ has mean 0, then $\|\mathbf{x}^*\| \leq \alpha^{-1/2}$ so the above results apply (see Fact M.1 in the Appendix). In fact when μ has mean 0, one will always have $\epsilon_d \geq \Omega(1)$ for $c_d = 1$ (as noted within the proof of Proposition 1.5). Proposition 1.5 represents the typical case in that μ can be “mildly off-centered”, for example by centering around a point \mathbf{x} drawn from μ itself (see again Fact M.1).

On the Smoothness of the Action Space in Online Optimization and Learning

Our positive results are in contrast with [Sel23, Proposition 3.9], which shows that $e^{\Omega(d)}$ time can be necessary for BIC exploration for r -regular μ . We believe that our results are not specific to the unit ball, but that the fundamental distinction between these two examples is the *smoothness* of the action set. In [Sel23, Proposition 3.9], the action set is a non-smooth polytope with corners, and the optimal action under the prior distribution is one of the corners. This means that given a small amount of new information, the posterior-optimal action will not change at all. By contrast our main algorithm crucially relies on expansiveness of the function

$$\ell^* \mapsto \arg \max_{\mathbf{x} \in \mathcal{A}} \langle \ell^*, \mathbf{x} \rangle.$$

In our setting \mathcal{A} is the unit ball and so this function is simply $\ell^* / \|\ell^*\|$. However we expect our methods to generalize to other smooth bodies, and plan to pursue this in future work.

It is worth mentioning that the geometry of the action set has long been understood to play an important role in high-dimensional learning and optimization. This was exemplified by the interplay between self-concordant barrier functions [NN94] and linear bandits via online stochastic mirror descent [AHR08, BCBK12, BCB⁺12, BCL18, BE19, KRdP21]. Geometric properties of the action space are also known to yield acceleration for full-information online learning and offline optimization [GH15, HLGS17, LK19, KdP21, Mha22, Mol23, TI24].

1.2 Additional Notation

We will use the following notation throughout the paper. Unless otherwise specified, $\|\cdot\|$ will refer to the ℓ_2 norm. We will use \mathbf{e}_i to refer to the i th vector of the standard basis. For a random variable X , we say that X is K -sub-gaussian if $\mathbb{P}(X \geq t) \leq 2e^{-t^2/K^2}$. We use the standard $f(d) = O(g(d))$ to mean that there exists some constant $C > 0$ such that $f(d) \leq C \cdot g(d)$ for all sufficiently large d . We also use $\mathcal{P}_S(\mathbf{u})$ to represent the projection of the vector \mathbf{u} onto the space S .

1.3 Outline

The rest of the paper will be organized as follows. In Section 2, we present sketches of our main algorithms and discuss the three key technical ideas behind the algorithm and analysis. In Section 3, we give the detailed main algorithm, present the two key technical propositions, and formally state our main theorem bounding the sample complexity of the algorithm.

2 Algorithm and Technical Overview

In this section, we present pseudocode for the main algorithm and give informal intuition for the key technical ideas used to show that this algorithm is BIC and satisfies the sample complexity bound of Theorem 1.3.

2.1 Algorithm Sketch

Before presenting the algorithm, we first state Lemma 2.1, which is the key observation used by the algorithm to select BIC actions. Informally, this lemma says that for any ψ that is a function of historical actions and rewards and possibly external randomness (but not on future information), the action $\mathbf{A}^{(t)}$ in the direction of $\mathbb{E}[\ell^* | \psi]$ is BIC (or any action is BIC if $\mathbb{E}[\ell^* | \psi] = \mathbf{0}$). Thus to prove an action $\mathbf{A}^{(t)}$ is BIC, it suffices to find such a ψ so that $\mathbf{A}^{(t)}$ is in the direction of $\mathbb{E}[\ell^* | \psi]$ whenever $\mathbb{E}[\ell^* | \psi] \neq \mathbf{0}$.

Lemma 2.1 (Proof in Appendix D). *Suppose that ψ is a function of the history before time t and potentially some external independent randomness ξ , i.e. $\psi = \psi(\mathbf{A}^{(1)}, r^{(1)}, \dots, \mathbf{A}^{(t-1)}, r^{(t-1)}, \xi)$. Let \mathbf{v} be any vector in \mathbb{R}^d . Define*

$$\text{Exploit}(\psi, \mathbf{v}) := \begin{cases} \frac{\mathbb{E}[\ell^* | \psi]}{\|\mathbb{E}[\ell^* | \psi]\|} & \text{if } \|\mathbb{E}[\ell^* | \psi]\| \neq \mathbf{0} \\ \mathbf{v} & \text{otherwise.} \end{cases}$$

Then $\mathbf{A}^{(t)} = \text{Exploit}(\psi, \mathbf{v})$ is BIC at time t .

Algorithm 1, together with Algorithms 2–3, presents a high-level sketch of the procedure used to prove Theorem 1.3. The algorithm first uses the prior-based BIC action (\mathbf{e}_1) for $\text{poly}(d)$ steps. In order to explore the rest of the action space, Algorithm 1 runs a single loop that repeatedly checks if λ -spectral exploration has been achieved. If λ -spectral exploration has not yet been achieved, then Algorithm 1 uses the subroutine InitialExploration to find a BIC action \mathbf{a} that has magnitude at least $\Omega(\epsilon_d)$ when projected onto the space of not-yet-sufficiently explored actions. Algorithm 1 then gives \mathbf{a} as input to the subroutine ExponentialGrowth. ExponentialGrowth returns another BIC action that has at least twice as large of a magnitude when projected onto the not-yet-sufficiently explored space of actions. Algorithm 1 repeatedly passes the action \mathbf{a} through ExponentialGrowth until the BIC action \mathbf{a} has magnitude of at least $\sqrt{\lambda}$ when projected onto the space of not-yet-sufficiently explored actions. Algorithm 1 then uses this action for $\text{poly}(d)$ steps. If λ -spectral exploration has still not been achieved, then Algorithm 1 explores a new direction by repeating the process of calling InitialExploration followed by repeated calls to ExponentialGrowth. In Sections 2.2 and 2.3, we discuss the main intuition of the subroutines InitialExploration and ExponentialGrowth respectively.

2.2 Initial Exploration

The goal of the InitialExploration routine (Algorithm 2) is to find a BIC action that has sufficient magnitude when projected onto the not-yet-sufficiently explored space of actions S^\perp . To do this, we first design an event E based on the historical actions and rewards H_t

Algorithm 1 BIC Exploration Pseudocode

- 1: Set $\mathbf{A}^{(t)} = \mathbf{e}_1$ for $\text{poly}(d)$ steps
 - 2: **while** λ -spectral exploration has not yet been achieved **do**
 - 3: $S \leftarrow$ space of actions that have already been sufficiently explored
 - 4: $\mathbf{a} \leftarrow \text{InitialExploration}(\cdot) \triangleright$ BIC-vector with $\Omega(\epsilon_d)$ -magnitude when projected onto S^\perp
 - 5: **while** magnitude of \mathbf{a} when projected onto S^\perp is less than $\sqrt{\lambda}$ **do**
 - 6: $\mathbf{a} \leftarrow \text{ExponentialGrowth}(\mathbf{a}) \triangleright$ new BIC-vector with double the magnitude when projected onto S^\perp
 - 7: **end while**
 - 8: Set $\mathbf{A}^{(t)} = \mathbf{a}$ for $\text{poly}(d)$ steps
 - 9: **end while**
-

Algorithm 2 InitialExploration Pseudocode

- 1: $\mathbf{M} \leftarrow \sum_{i=1}^{t-1} (\mathbf{A}^{(i)})^{\otimes 2}$
 - 2: $\mathbf{w}_1, \dots, \mathbf{w}_{\ell_\lambda} \leftarrow$ orthonormal eigenvectors of \mathbf{M} with corresponding eigenvalues greater than λ
 - 3: $S \leftarrow \text{Span}(\mathbf{w}_1, \dots, \mathbf{w}_{\ell_\lambda}) \triangleright$ Space of already-sufficiently explored actions
 - 4: $\hat{y}_i \leftarrow$ empirical estimate of $\langle \ell^*, \mathbf{w}_i \rangle$ for $i \leq \ell_\lambda$
 - 5: $\mathbf{z} \leftarrow \mathbb{E}[\ell^* \mid \hat{\mathbf{y}}]$
 - 6: $f \leftarrow$ function of \mathbf{z} such that $\mathbb{E}[\mathbf{z}f(\mathbf{z})] = 0$ and $f(\mathbf{z}) \in [\Omega(\epsilon_d), 1]$.
 - 7: $E \leftarrow \{\text{Bernoulli}(f(\mathbf{z})) = 1\}$
 - 8: Set $\mathbf{A}^{(t)} = \text{Exploit}(1_E, \mathbf{v})$ for $\mathbf{v} \in S^\perp \triangleright \mathbf{A}^{(t)}$ explores a new direction with probability $f(\mathbf{z})$
 - 9: $r \leftarrow r^{(t)}$ if we explored and otherwise $r \leftarrow N(0, 1)$
 - 10: **return** $\text{Exploit}(\text{sign}(r), \mathbf{v}) \triangleright$ where \mathbf{v} is in S^\perp
-

Algorithm 3 ExponentialGrowth Pseudocode

- 1: $S \leftarrow$ space of actions that have already been sufficiently explored
 - 2: Set $\mathbf{A}^{(t)} = \mathbf{a}$ for $\text{poly}(d)$ steps to observe a set of rewards $\{r^{(t)}\}$.
 - 3: $R \leftarrow$ average of rewards from the component of \mathbf{a} projected onto the space of not-yet-sufficiently explored actions (S^\perp)
 - 4: **return** $\text{Exploit}(\text{sign}(R), \mathbf{v}) \triangleright$ where \mathbf{v} is in S^\perp
-

and external randomness such that $\mathbb{P}(E \mid H_t) \geq \Omega(\epsilon_d)$ for all histories H_t and such that conditional on E , ℓ^* has 0 expectation when projected onto any direction in S . We then choose the BIC action $A^{(t)}$ in Line 8 so that $A^{(t)}$ explores S^\perp whenever event E holds. More concretely, the second condition on E implies that the action $A^{(t)} = \text{Exploit}(1_E, \mathbf{v})$ for any $\mathbf{v} \in S^\perp$ will satisfy $A^{(t)} \in S^\perp$ whenever event E holds, and $A^{(t)}$ is BIC by Lemma 2.1. Using this BIC action on Line 8 therefore explores S^\perp whenever event E holds. We define the signal r as the reward at time t if event E holds and otherwise as independent $N(0, 1)$ noise. We next show that, conditional on the sign of r , the expectation of ℓ^* always has sufficient magnitude when projected onto S^\perp (see Lemma A.1). This implies that for any $\mathbf{v} \in S^\perp$, the action $\text{Exploit}(\text{sign}(r), \mathbf{v})$ will have sufficient magnitude when projected onto S^\perp as desired. Therefore, we can return the action $\text{Exploit}(\psi, \mathbf{v})$, which is BIC by Lemma 2.1.

All that remains is to define the event E from the previous paragraph. Formally, we want an event E that is a function of the historical actions and rewards and independent random variable ξ such that $\mathbb{E}[\langle \ell^*, \mathbf{w}_i \rangle \mid E] = 0$ for all $i \leq \ell_\lambda$ and such that $\mathbb{P}(E \mid H_t) \geq \Omega(\epsilon_d)$ for all histories H_t . The key to constructing E is Lemma 2.2, which implies that for any random variable \mathbf{x} not confined to any half-spaces, there exists a function f such that \mathbf{x} has expectation equal to $\mathbf{0}$ conditional on the event $\{\text{Bernoulli}(f(\mathbf{x})) = 1\}$

Lemma 2.2 (Proof in Appendix G). *Let μ be a probability distribution on \mathbb{R}^d with finite first moment and suppose for $0 < \epsilon \leq 1/2$ we have that*

$$\min_{\|\mathbf{v}\|=1} \mathbb{E}^{\mathbf{x} \sim \mu}[\langle \mathbf{v}, \mathbf{x} \rangle_+] \geq \epsilon.$$

Then there exists a Borel measurable function $f : \mathbb{R}^d \rightarrow \left[\frac{\epsilon}{4 \max(\|\mathbb{E}[\mathbf{x}]\|, 1)}, 1 \right]$ with $\mathbb{E}[\mathbf{x}f(\mathbf{x})] = \mathbf{0}$.

If we knew the exact values of the vector $\mathbf{x} := (\langle \ell^*, \mathbf{w}_1 \rangle, \dots, \langle \ell^*, \mathbf{w}_{\ell_\lambda} \rangle)$, then we could directly apply Lemma 2.2 to \mathbf{x} and use the resulting function f to define the event $E = \{\text{Bernoulli}(f(\mathbf{x})) = 1\}$. This event E would satisfy the desired property that $\mathbb{E}[\langle \ell^*, \mathbf{w}_i \rangle \mid E] = 0$ for $i \leq \ell_\lambda$ and that $\Pr(E \mid H_t) = \Omega(\epsilon_d)$ for all H_t . However, we do not know the exact values of $\langle \ell^*, \mathbf{w}_1 \rangle, \dots, \langle \ell^*, \mathbf{w}_{\ell_\lambda} \rangle$ because we do not know ℓ^* , and therefore this event E is not a function of historical actions and rewards. Instead, we can estimate $\langle \ell^*, \mathbf{w}_i \rangle$ as \hat{y}_i using the historical actions and returns. These estimates will be relatively accurate because these directions are already well-explored. Defining $\mathbf{z} = \mathbb{E}[\ell^* \mid \hat{\mathbf{y}}]$, we show that \mathbf{z} also satisfies the assumption of Lemma 2.2 (Lemma A.4). Therefore, applying Lemma 2.2 to \mathbf{z} gives a function f such that $f(\mathbf{z}) \geq \Omega(\epsilon_d)$ for all \mathbf{z} . \mathbf{z} is a function of historical actions and rewards and external randomness as desired. Defining the event $E = \{\text{Bernoulli}(f(\mathbf{z})) = 1\}$, we have as desired that $\mathbb{E}[\langle \ell^*, \mathbf{w}_i \rangle \mid E] = 0$ for all $i \leq \ell_\lambda$ and that $\mathbb{P}(E \mid H_t) = \Omega(\epsilon_d)$ for all H_t .

2.3 Exponential Growth

The goal of the ExponentialGrowth routine (Algorithm 3) is to take a BIC action \mathbf{a} and return a new BIC action that has twice as large of a magnitude when projected onto the not-yet-sufficiently explored space of actions S^\perp . Using Lemma 2.1, we will find a signal R such that for any $\mathbf{v} \in S^\perp$, the action $\text{Exploit}(R, \mathbf{v})$ will have twice as large magnitude when projected onto S^\perp as \mathbf{a} has.

The key intuition is that because the action space is curved, conditioning on noisy information about the sign of ℓ^* in a specific direction will increase the magnitude of the expectation of ℓ^* projected in that direction. Consider the following simplified example. Suppose that $d = 2$. Also suppose we already know ℓ^*_1 , and the goal is to explore ℓ^*_2 . Furthermore, assume that the initial BIC action is $\mathbf{A}^{(t)}$ shown in the left-most diagram of Figure 1 that has ϵ as the y -coordinate. Using this action gives a rewards $r^{(t)}$. Because we know the value of ℓ^*_1 , we can remove this component of the reward $r^{(t)}$, and we are left with a signal $r = \epsilon \ell^*_2 + N(0, \sigma^2)$ for some $\sigma^2 > 0$. If we take $\mathbf{A}^{(t+1)}$ to be the unit vector in the direction of $\mathbb{E}[\ell^* | r]$, then we will have that the new y -coordinate is 2ϵ . By Lemma 2.1, this choice of $\mathbf{A}^{(t+1)}$ is BIC, which is an important consequence of the curved action space. We then observe $r^{(t+1)}$, and we can repeat this process again to find a BIC action $\mathbf{A}^{(t+2)}$ that has y -coordinate of 4ϵ . Therefore, in this simplified example we are able to exponentially grow the y -coordinate for BIC actions, which is a consequence of the curvature of the action space. This process is demonstrated in Figure 1.

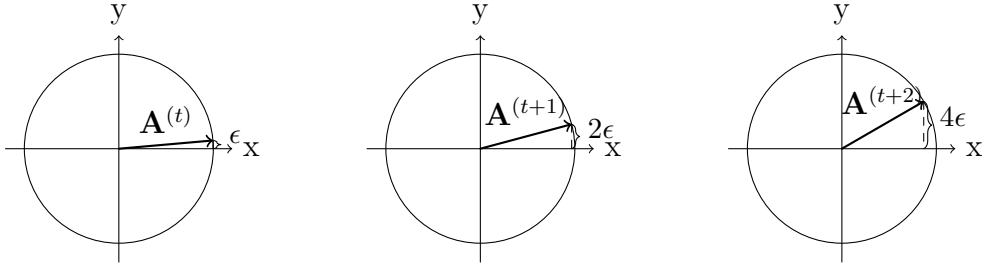


Figure 1: Diagram illustrating exponentially growing exploration. First, we have a BIC action $\mathbf{A}^{(t)}$ with y -coordinate ϵ . Using $r^{(t)}$, we design a signal conditional on which the expectation of ℓ^*_2 doubles. Exploit(\cdot) then gives a BIC action $\mathbf{A}^{(t+1)}$ with y -coordinate 2ϵ . Using $r^{(t+1)}$, we again double the conditional expectation of ℓ^*_2 and get BIC action $\mathbf{A}^{(t+2)}$ with y -coordinate 4ϵ . Increasing the conditional expectation of ℓ^*_2 gives a BIC action with a larger y -coordinate, and this action’s feedback gives a stronger signal that more rapidly increases the conditional expectation of ℓ^*_2 . This allows us to “bootstrap” an exponentially weak starting signal all the way to constant signal strength without suffering exponential sample complexity.

Equipped with the intuition of the previous paragraph, we now analyze Algorithm 3. Recall that S is the space of actions that have already been sufficiently explored. Algorithm 3 first uses the action $\mathbf{A}^{(t)} = \mathbf{a}$ for $\text{poly}(d)$ steps. Taking an average of the rewards $\{r^{(t)}\}$ from these actions gives a close estimate of $\langle \mathbf{a}, \ell^* \rangle$. Using the previous observed actions and rewards, we can remove the component of $\langle \mathbf{a}, \ell^* \rangle$ that comes from \mathbf{a} projected onto S . This leaves a signal R which is the average of rewards from the component of \mathbf{a} projected onto just S^\perp . Using concentration laws for conditional probabilities (Lemmas A.2 and A.3), we formalize the intuition from the previous paragraph to show that for any $\mathbf{v} \in S^\perp$, the projection of $\text{Exploit}(R, \mathbf{v})$ onto S^\perp will always have magnitude at least $2\mathcal{P}_{S^\perp}(\mathbf{a})$. The action $\text{Exploit}(R, \mathbf{v})$ is BIC by Lemma 2.1. Therefore, we have found a new BIC action $\text{Exploit}(R, \mathbf{v})$ that has twice as large magnitude when projected onto S^\perp as \mathbf{a} .

2.4 Pushing eigenvalues upwards

An important step of the proof of Algorithm 1 is showing that the while loop will terminate in polynomial time. To show this, we show that the action \mathbf{a} used for $\text{poly}(d)$ steps at the end of each round of the while loop sufficiently increases the eigenvalues of $\mathbf{M}^{(t)} := \sum_{i=1}^t (\mathbf{A}^{(i)})^{\otimes 2}$. Note that at each time t , the matrix $\mathbf{M}^{(t)}$ increases by a rank-1 update, i.e. $\mathbf{M}^{(t+1)} = \mathbf{M}^{(t)} + (\mathbf{A}^{(t+1)})^{\otimes 2}$. In order to show that we will eventually achieve λ -spectral exploration, we must show that the small eigenvalues of $\mathbf{M}^{(t+1)}$ increase relative to the small eigenvalues of $\mathbf{M}^{(t)}$ after each round of the while loop. Although we do not follow this route, we mention that [Gol73] provides exact descriptions for the eigenvalues of rank-one updates as above, giving a rational function $\omega(x)$ (with coefficients depending on $\mathbf{M}^{(t)}$ and $\mathbf{A}^{(t+1)}$) which has roots equal to the eigenvalues of $\mathbf{M}^{(t+1)}$. However, it is not clear how helpful this is for the quantitative estimates we require.

At first glance, we might hope that we can sufficiently increase all of the eigenvalues of $\mathbf{M}^{(t)}$ in just d rounds if in each round we partially explore a new not-yet-explored direction. However, this unfortunately is not always the case. For example, suppose in the first round we use BIC action $\mathbf{A}^{(1)} = (1, 0, 0, \dots, 0)$. Writing $\varphi = 1/\sqrt{5}$, we then in the next $d-1$ actions use the BIC actions $\mathbf{A}^{(2)} = (2\varphi, -\varphi, 0, 0, \dots, 0)$, $\mathbf{A}^{(3)} = (0, 2\varphi, -\varphi, 0, 0, \dots, 0)$, and so on until $\mathbf{A}^{(d)} = (0, 0, \dots, 0, 2\varphi, -\varphi)$. Then each $\mathbf{A}^{(i)}$ has distance $\varphi = \Omega(1)$ from the span of the preceding actions, so we might hope that this already yields $\Omega(1)$ -spectral exploration. However, note that the vector $\mathbf{x} = (1, 2, 4, \dots, 2^{d-1})$ satisfies $\langle \mathbf{x}, \mathbf{A}^{(i)} \rangle = 1_{i=1} \leq 2^{-(d-1)} \|\mathbf{x}\|$ for all $1 \leq i \leq d$. It follows that the matrix $\mathbf{M}^{(d)} = \sum_{i=1}^d (\mathbf{A}^{(i)})^{\otimes 2}$ has smallest eigenvalue which is *exponentially small* in d (since $\langle \mathbf{x}, \mathbf{M}^{(d)} \mathbf{x} \rangle / \|\mathbf{x}\|^2$ is exponentially small).

We show that our algorithm terminates in $\text{poly}(d)$ steps using the linear algebraic Lemma 2.3 below, which may be of independent interest. Informally, this lemma says that if \mathbf{u} has non-negligible projection onto the space orthogonal to the large-or-medium eigenvalues of \mathbf{M} , then the rank-1 update of $\mathbf{M} + \mathbf{u}^{\otimes 2}$ increases the total sum of the small-or-medium eigenvalues non-negligibly.

Lemma 2.3 (Proof in Appendix E). *Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j \in \mathbb{R}^d$ such that $\|\mathbf{v}_i\| = 1$. Define $\mathbf{M} = \sum_{i=1}^j \mathbf{v}_i^{\otimes 2}$. Suppose $\mathbf{w}_1, \dots, \mathbf{w}_d$ are orthonormal eigenvectors of \mathbf{M} with corresponding eigenvalues $\lambda_1 \geq \dots \geq \lambda_d \geq 0$. Define ℓ_ϵ as the largest index such that $\lambda_j \geq \epsilon$ for some $0 < \epsilon < 1$, and define $S = \text{Span}(\mathbf{w}_1, \dots, \mathbf{w}_{\ell_\epsilon})$. Suppose \mathbf{u} is a vector such that $\|P_{S^\perp}(\mathbf{u})\|_2^2 \geq \epsilon$, and define $\mathbf{M}' = \mathbf{M} + \mathbf{u}^{\otimes 2}$. Let $\mathbf{w}'_1, \dots, \mathbf{w}'_d$ be the orthonormal eigenvectors of \mathbf{M}' with corresponding eigenvalues $\lambda'_1 \geq \dots \geq \lambda'_d$. Finally, let ℓ be the largest index such that $\lambda_\ell \geq \frac{200d^3}{\epsilon^2}$. Then*

$$\sum_{i=\ell+1}^d \lambda'_i \geq \epsilon/2 + \sum_{i=\ell+1}^d \lambda_i.$$

3 Main Results

In this section, we will formally state our main algorithms and theorems. For presentation purposes, for the formal algorithms and theorems we will define

$$\lambda := \min \left(1, \min(\delta_{LA.2}, \delta_{LA.3}, 1/c_{LA.3})^2 \frac{(c_v / \sqrt{8\pi})^2}{4d(K\sqrt{\pi} + 1)^2} \right) = \Omega(c_v / d),$$

where $\delta_{LA.2}$ is a constant from Lemma A.2 and $\delta_{LA.3}, c_{LA.3}$ are constants from Lemma A.3. Note that if we can do λ -spectral exploration in T rounds with this value of λ , then for any $\bar{\lambda}$ we can do $\bar{\lambda}$ -spectral exploration in $\lceil \frac{\bar{\lambda}}{\lambda} \rceil \cdot T$ rounds by simply repeating the algorithm $\lceil \frac{\bar{\lambda}}{\lambda} \rceil$ times. Therefore, if we can ensure λ -spectral exploration for the λ value described above in $\text{poly}(d)$ rounds, then we can also ensure $\bar{\lambda}$ -spectral exploration in $\bar{\lambda} \cdot \text{poly}(d)$ rounds for any $\bar{\lambda} \geq \lambda$.

3.1 Algorithms

Our main algorithm is presented in Algorithm 4, with subroutines in Algorithms 5 and 6. Note that these three algorithms directly correspond to the pseudocode presented in Algorithms 1–3.

Lemma 3.1 shows that the application of Lemma 2.2 (in the form of Lemma A.4) in Algorithm 5 is valid.

Lemma 3.1 (Proof in App C). *Every time Algorithm 5 (Line 2) calls Algorithm 4 to define \hat{y}_ℓ , we have $\hat{y}_\ell \stackrel{d}{=} x_\ell^* + N(0, c_{LA.4} \frac{\lambda}{\lambda_\ell})$. Thus, $\mathbf{z}(\hat{\mathbf{y}})$ (Line 4) satisfies the assumptions of Lemma 2.2 with $\epsilon = \frac{\epsilon_d c_d}{4}$.*

3.2 Propositions and Theorem

As discussed in Section 2.2, the main purpose of Algorithm 5 is to find a BIC action \mathbf{a} that has sufficiently high magnitude when projected onto the space of not-yet-sufficiently explored actions. This is formalized in Proposition 3.2. As discussed in Section 2.3, the main purpose of Algorithm 6 is to double the magnitude of \mathbf{a} when projected onto the space of not-yet-sufficiently explored actions. This is formalized in Proposition 3.3.

Proposition 3.2 (Proof in Appendix I). *The action \mathbf{a} returned by Algorithm 5 satisfies*

$$\|\mathcal{P}_{S^\perp}(\mathbf{a})\| \geq \frac{c_{LA.1} c_v^{2.5} \epsilon_d c_d}{16 (K\sqrt{\pi} + 1)} := c_{P3.2} c_v^{2.5} \epsilon_d c_d.$$

Proposition 3.3 (Proof in Appendix J). *The action returned by Algorithm 6 satisfies*

$$\|\mathcal{P}_{S^\perp}(\mathbf{b})\| \geq 2 \|\mathcal{P}_{S^\perp}(\mathbf{a})\|.$$

We can now state our main theorem bounding the sample complexity of Algorithm 4.

Algorithm 4 BIC Exploration

Input: λ

```

1:  $\kappa \leftarrow \max\left(\frac{1}{\lambda_{cL\text{A.4}}}, \frac{4d(K\sqrt{\pi}+1)(1+\frac{1}{\lambda})}{c_v^2/(8\pi)}\right)$ 
2:  $\mathbf{v}_1 \leftarrow \mathbf{e}_1$ 
3: for  $t \in [0 : \kappa)$  do
4:   Set  $\mathbf{A}^{(t)} = \mathbf{v}_1$ 
5:    $q_1^t \leftarrow r^{(t)}$ 
6: end for
7:  $t \leftarrow \kappa$ 
8:  $j \leftarrow 1$ 
9: while minimum eigenvalue of  $\mathbf{M} = \sum_{i=1}^j \mathbf{v}_i^{\otimes 2}$  is smaller than  $\lambda$  do
10:   $\mathbf{w}_1, \dots, \mathbf{w}_d \leftarrow$  orthonormal eigenvectors of  $\mathbf{M}$  with corresponding eigenvalues  $\lambda_1 \geq \dots \geq \lambda_d$ .
11:   $\ell_\lambda \leftarrow \max\{i : \lambda_i \geq \lambda\}$ 
12:   $S \leftarrow \text{Span}(\mathbf{w}_1, \dots, \mathbf{w}_{\ell_\lambda})$ 
13:   $\mathbf{a}, t \leftarrow \text{InitialExploration}(\{\mathbf{w}_i\}_{i=1}^d, \{\lambda_i\}_{i=1}^{\ell_\lambda}, \{\mathbf{v}_i\}_{i=1}^j, \{\{q_i^{t'}\}_{t'=0}^{\kappa-1}\}_{i=1}^j, t)$ 
14:  while  $\|\mathcal{P}_{S^\perp}(\mathbf{a})\| \leq \sqrt{\lambda}$  do
15:     $\mathbf{a}, t \leftarrow \text{ExponentialGrowth}(\mathbf{a}, \{\mathbf{w}_i\}_{i=1}^{\ell_\lambda}, \{\lambda_i\}_{i=1}^{\ell_\lambda}, \{\mathbf{v}_i\}_{i=1}^j, \{\{q_i^{t'}\}_{t'=0}^{\kappa-1}\}_{i=1}^j, t)$ 
16:  end while
17:   $\mathbf{v}_{j+1} \leftarrow \mathbf{a}$ 
18:  for  $t' \in [0 : \kappa)$  do
19:    Set  $\mathbf{A}^{(t+t')} = \mathbf{a}$ 
20:     $q_{j+1}^{t'} \leftarrow r^{(t+t')}$ 
21:  end for
22:   $t \leftarrow t + \kappa$ 
23:   $j \leftarrow j + 1$ 
24: end while

```

Algorithm 5 InitialExploration

Input: $\{\mathbf{w}_i\}_{i=1}^d, \{\lambda_i\}_{i=1}^{\ell_\lambda}, \{\mathbf{v}_i\}_{i=1}^j, \{\{q_i^{t'}\}_{t'=0}^{\kappa-1}\}_{i=1}^j, t$

- 1: Define $\mathbf{x}^* \in \mathbb{R}^{\ell_\lambda}$ as $x_\ell^* = \langle \ell^*, \mathbf{w}_\ell \rangle$.
 - 2: Define $\hat{\mathbf{y}} \in \mathbb{R}^{\ell_\lambda}$ as $\hat{y}_\ell = \lambda c_{L\text{A.4}} \sum_{t'=0}^{\frac{1}{\lambda c_{L\text{A.4}}}-1} \sum_{k=1}^j \frac{\langle \mathbf{v}_k, \mathbf{w}_\ell \rangle}{\lambda_\ell} q_k^{t'}$ $\triangleright \hat{y}_\ell \sim x_\ell^* + N(0, c_{L\text{A.4}} \frac{\lambda}{\lambda_\ell})$ by Lemma A.5 via Lemma 3.1
 - 3: $\mathbf{z}(\mathbf{y}) \leftarrow \mathbb{E}[\mathbf{x}^* \mid \hat{\mathbf{y}} = \mathbf{y}]$
 - 4: $f \leftarrow$ function from Lemma 2.2 for $\mathbf{z}(\hat{\mathbf{y}})$ with $\epsilon = \frac{\epsilon_d c_d}{4}$. $\triangleright f$ exists by Lemma A.4 via Lemma 3.1
 - 5: $\Psi \leftarrow \text{Bernoulli}(f(\mathbf{z}(\hat{\mathbf{y}})))$
 - 6: Set $\mathbf{A}^{(t)} = \text{Exploit}(\Psi, \mathbf{w}_{\ell_\lambda+1})$
 - 7: $R \leftarrow \begin{cases} r^{(t)} & \text{w.p. } \frac{\epsilon_d c_d}{16(K\sqrt{\pi}+1)f(\mathbf{z}(\hat{\mathbf{y}}))} \text{ if } \Psi = 1 \\ N(0, 1) & \text{otherwise} \end{cases}$ \triangleright The above equation involves valid probabilities by Lemma 2.2 and because $\max(\|\mathbb{E}[\mathbf{z}(\hat{\mathbf{y}})]\|, 1) \leq \|\mathbb{E}[\mathbf{x}^*]\| + 1 \leq K\sqrt{\pi} + 1$
 - 8: $\mathbf{a} \leftarrow \text{Exploit}(1_{R>0}, \mathbf{w}_{\ell_\lambda+1})$.
 - 9: **return** $\mathbf{a}, t + 1$
-

Algorithm 6 ExponentialGrowth

Input: $\mathbf{a}, \{\mathbf{w}_i\}_{i=1}^{\ell_\lambda}, \{\lambda_i\}_{i=1}^{\ell_\lambda}, \{\mathbf{v}_i\}_{i=1}^j, \{\{q_i^{t'}\}_{t'=0}^{\kappa-1}\}_{i=1}^j, t$

- 1: $S \leftarrow \text{Span}(\mathbf{w}_1, \dots, \mathbf{w}_{\ell_\lambda})$
 - 2: $\mathbf{x} \leftarrow \frac{\mathcal{P}_{S^\perp}(\mathbf{a})}{\|\mathcal{P}_{S^\perp}(\mathbf{a})\|}$
 - 3: $c_k \leftarrow \sum_{i=1}^{\ell_\lambda} \frac{\langle \mathcal{P}_S(\mathbf{a}), \mathbf{w}_i \rangle \langle \mathbf{v}_k, \mathbf{w}_i \rangle}{\lambda_i}$ for $k \in [1 : j]$ $\triangleright \mathcal{P}_S(\mathbf{a}) = \sum_{k=1}^j c_k \mathbf{v}_k$ by Lemma A.5
 - 4: $L \leftarrow \frac{4d(\mathbb{E}[\ell^*]+1)^2(1+\sum_{k=1}^j c_k^2)}{c_{L\text{A.2}}^2}$
 - 5: For $t' \in [t, t+L)$, set $\mathbf{A}^{(t')} = \mathbf{a}$
 - 6: $t \leftarrow t + L$
 - 7: $R \leftarrow \sum_{t'=t}^{t+L-1} \left(r^{(t')} - \sum_{k=1}^j c_k q_k^{t'} \right)$
 - 8: $\mathbf{b} \leftarrow \text{Exploit}(1_{R>0}, \mathbf{w}_{\ell_\lambda+1})$
 - 9: **return** \mathbf{b}, t
-

Theorem 3.4 (Proof in Appendix K). *Algorithm 4 is BIC and has sample complexity*

$$O\left(\log\left(\frac{1}{c_v \epsilon_d c_d}\right)\left(\frac{d^5}{\lambda^4 c_v^2} + \frac{d^4}{c_d^2 \lambda^4}\right)\right).$$

As discussed above, for any $\bar{\lambda}$, we can repeat Algorithm 4 for $\lceil \bar{\lambda}/\lambda \rceil$ times to get $\bar{\lambda}$ -spectral exploration. This is because trace is additive, and therefore if running Algorithm 4 once gives λ -spectral exploration, then running it $\lceil \bar{\lambda}/\lambda \rceil$ times will give $\bar{\lambda}$ -spectral exploration. Multiplying the bound from Theorem 3.4 by $\lceil \bar{\lambda}/\lambda \rceil$ gives that $\bar{\lambda}$ -spectral exploration is achievable in $\bar{\lambda} \left(\frac{d}{c_v + c_d}\right)^{O(1)} \log(1/\epsilon_d)$ rounds, matching the desired result of Theorem 1.3.

4 Discussion

We conclude with a more detailed discussion on how our results can be combined with the results of [Sel23] to achieve end-to-end guarantees for incentivized exploration. Here we focus on r -regular μ as assumed in that work, which is encapsulated by Proposition 1.4. Recall [Sel23, Theorem 3.5] shows that for $\epsilon > 0$, if an algorithm has already achieved $\tilde{O}(d^4/r^2\epsilon^2)$ -spectral exploration at time t , then running Thompson sampling from time t onward will be ϵ -BIC (where ϵ -BIC relaxes Definition 1.1 by subtracting an ϵ term from the right-hand side).

Our main result (Theorem 1.3) efficiently achieves the necessary spectral exploration, with at most $\text{poly}(d, 1/r, 1/\epsilon)$ sample complexity (and thus additional regret). Note that our algorithm actually gives a stronger guarantee than in [Sel23] (BIC rather than ϵ -BIC). If we only need to guarantee the initial exploration is ϵ -BIC, then we no longer need the InitialExploration phase of the algorithm, and therefore can drop Assumption 1. Combining our result with [Sel23] and the analysis of Thompson sampling in [DVR18] or [AG13], we therefore obtain an end-to-end ϵ -BIC algorithm with respectively Bayesian regret $\text{poly}(d, 1/r, 1/\epsilon) + \tilde{O}(d\sqrt{T})$ or frequentist regret $\text{poly}(d, 1/r, 1/\epsilon) + \tilde{O}(d^{3/2}\sqrt{T})$. Namely, one first runs our algorithm for $\text{poly}(d, 1/r, 1/\epsilon)$ steps to guarantee the required spectral exploration, and then uses Thompson Sampling for all remaining steps. Since the regret from the initial exploration phase is constant relative to T (and polynomial in d), this combined algorithm will asymptotically obey the state-of-the-art regret bounds for Thompson Sampling.

References

- [AG13] S. Agrawal and N. Goyal. Thompson sampling for contextual bandits with linear payoffs. In *International Conference on Machine Learning*, pages 127–135. PMLR, 2013.
- [AG17] S. Agrawal and N. Goyal. Near-optimal regret bounds for Thompson Sampling. *Journal of the ACM (JACM)*, 64(5):1–24, 2017.
- [AGZ10] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An Introduction to Random Matrices*. Number 118. Cambridge university press, 2010.
- [AHR08] J. D. Abernethy, E. Hazan, and A. Rakhlin. Competing in the Dark: An Efficient Algorithm for Bandit Linear Optimization. In *COLT*, pages 263–274. Citeseer, 2008.

- [AT20] P. Agrawal and T. Tulabandhula. Incentivising exploration and recommendations for contextual bandits with payments. In *Multi-Agent Systems and Agreement Technologies: 17th European Conference, EUMAS 2020, and 7th International Conference, AT 2020, Thessaloniki, Greece, September 14-15, 2020, Revised Selected Papers 17*, pages 159–170. Springer, 2020.
- [BCB⁺12] S. Bubeck, N. Cesa-Bianchi, et al. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5(1):1–122, 2012.
- [BCBK12] S. Bubeck, N. Cesa-Bianchi, and S. M. Kakade. Towards minimax policies for online linear optimization with bandit feedback. In *Conference on Learning Theory*, pages 41–1. JMLR Workshop and Conference Proceedings, 2012.
- [BCL18] S. Bubeck, M. Cohen, and Y. Li. Sparsity, variance and curvature in multi-armed bandits. In *Algorithmic Learning Theory*, pages 111–127. PMLR, 2018.
- [BE19] S. Bubeck and R. Eldan. The entropic barrier: Exponential families, log-concave geometry, and self-concordance. *Mathematics of Operations Research*, 44(1):264–276, 2019.
- [BL13] S. Bubeck and C.-Y. Liu. Prior-free and prior-dependent regret bounds for Thompson Sampling. *Advances in neural information processing systems*, 26, 2013.
- [BM19] D. Bergemann and S. Morris. Information Design: A Unified Perspective. *Journal of Economic Literature*, 57(1):44–95, 2019.
- [BS22] S. Bubeck and M. Sellke. First-Order Bayesian Regret Analysis of Thompson Sampling. *IEEE Transactions on Information Theory*, 69(3):1795–1823, 2022.
- [CH18] Y.-K. Che and J. Hörner. Recommender systems as mechanisms for social learning. *The Quarterly Journal of Economics*, 133(2):871–925, 2018.
- [CP23] S. Chewi and A.-A. Pooladian. An entropic generalization of Caffarelli’s contraction theorem via covariance inequalities. *Comptes Rendus. Mathématique*, 361(G9):1471–1482, 2023.
- [DCWY19] R. Dwivedi, Y. Chen, M. J. Wainwright, and B. Yu. Log-concave sampling: Metropolis-Hastings algorithms are fast. *Journal of Machine Learning Research*, 20(183):1–42, 2019.
- [DM19] A. Durmus and É. Moulines. High-dimensional Bayesian inference via the Unadjusted Langevin Algorithm. *Bernoulli*, 25(4A), 2019.
- [DVR18] S. Dong and B. Van Roy. An Information-Theoretic Analysis for Thompson Sampling with Many Actions. *Advances in Neural Information Processing Systems*, 31, 2018.
- [FKKK14] P. Frazier, D. Kempe, J. Kleinberg, and R. Kleinberg. Incentivizing exploration. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 5–22, 2014.
- [GH15] D. Garber and E. Hazan. Faster rates for the Frank-Wolfe method over strongly-convex sets. In *International Conference on Machine Learning*, pages 541–549. PMLR, 2015.
- [GKS24] Y. Gur, A. Kalvit, and A. Slivkins. Incentivized Exploration via Filtered Posterior Sampling. In *Proceedings of the 25th ACM Conference on Economics and Computation*, pages 1200–1200, 2024.
- [Gol73] G. H. Golub. Some modified matrix eigenvalue problems. *SIAM review*, 15(2):318–334, 1973.
- [Gor41] R. D. Gordon. Values of Mills’ ratio of area to bounding ordinate and of the normal probability integral for large values of the argument. *Ann. Math. Stat.*, 12(3):364–366, 1941.
- [HLGS17] R. Huang, T. Lattimore, A. György, and C. Szepesvári. Following the leader and fast rates in online linear prediction: Curved constraint sets and other regularities. *Journal of Machine Learning Research*, 18(145):1–31, 2017.
- [HNSW22] X. Hu, D. D. Ngo, A. Slivkins, and Z. S. Wu. Incentivizing Combinatorial Bandit Exploration. *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

- [IEV23] N. Immorlica, F. Echenique, and V. V. Vazirani. *Online and Matching-Based Market Design*. Cambridge University Press, 2023.
- [IMSW19] N. Immorlica, J. Mao, A. Slivkins, and Z. S. Wu. Bayesian exploration with heterogeneous agents. In *The world wide web conference*, pages 751–761, 2019.
- [IMSW20] N. Immorlica, J. Mao, A. Slivkins, and Z. S. Wu. Incentivizing Exploration with Selective Data Disclosure. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 647–648, 2020.
- [Kam19] E. Kamenica. Bayesian Persuasion and Information Design. *Annual Review of Economics*, 11:249–272, 2019.
- [KdP21] T. Kerdreux, A. d’Aspremont, and S. Pokutta. Projection-free optimization on uniformly convex sets. In *International Conference on Artificial Intelligence and Statistics*, pages 19–27. PMLR, 2021.
- [KKM12] E. Kaufmann, N. Korda, and R. Munos. Thompson sampling: An asymptotically optimal finite-time analysis. In *International conference on algorithmic learning theory*, pages 199–213. Springer, 2012.
- [KKM⁺17] S. Kannan, M. Kearns, J. Morgenstern, M. Pai, A. Roth, R. Vohra, and Z. S. Wu. Fairness incentives for myopic agents. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 369–386, 2017.
- [KMP14] I. Kremer, Y. Mansour, and M. Perry. Implementing the “Wisdom of the Crowd”. *Journal of Political Economy*, 122(5):988–1012, 2014.
- [KRdP21] T. Kerdreux, C. Roux, A. d’Aspremont, and S. Pokutta. Linear bandits on uniformly convex sets. *Journal of Machine Learning Research*, 22(284):1–23, 2021.
- [LK19] K. Levy and A. Krause. Projection Free Online Learning over Smooth Sets. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 1458–1466. PMLR, 2019.
- [Mha22] Z. Mhammedi. Exploiting the curvature of feasible sets for faster projection-free online learning. *arXiv preprint arXiv:2205.11470*, 2022.
- [Mol23] M. Molinaro. Strong Convexity of Feasible Sets in Off-line and Online Optimization. *Mathematics of Operations Research*, 48(2):865–884, 2023.
- [MSS20] Y. Mansour, A. Slivkins, and V. Syrgkanis. Bayesian Incentive-Compatible Bandit Exploration. *Operations Research*, 68(4):1132–1161, 2020. Preliminary version in *ACM EC 2015*.
- [MSSW22] Y. Mansour, A. Slivkins, V. Syrgkanis, and S. Wu. Bayesian exploration: Incentivizing exploration in Bayesian games. *Operations Research*, 70(2), 2022. Preliminary version in *EC 2016*.
- [Mye86] R. B. Myerson. Multistage games with communication. *Econometrica: Journal of the Econometric Society*, pages 323–358, 1986.
- [NN94] Y. Nesterov and A. Nemirovskii. *Interior-point polynomial algorithms in convex programming*. SIAM, 1994.
- [RV14] D. Russo and B. Van Roy. Learning to Optimize via Posterior Sampling. *Mathematics of Operations Research*, 39(4):1221–1243, 2014.
- [RV16] D. Russo and B. Van Roy. An information-theoretic analysis of thompson sampling. *JMLR*, 17:68:1–68:30, 2016.
- [Sel23] M. Sellke. Incentivizing exploration with linear contexts and combinatorial actions. In *International Conference on Machine Learning*, pages 30570–30583. PMLR, 2023.
- [Sli19] A. Slivkins. Introduction to multi-armed bandits. *Foundations and Trends® in Machine Learning*, 12(1-2):1–286, November 2019.

- [SS23] M. Sellke and A. Slivkins. The Price of Incentivizing Exploration: A Characterization via Thompson Sampling and Sample Complexity. *Operations Research*, 71(5):1706–1732, 2023.
- [SS24] M. Simchowitz and A. Slivkins. Exploration and incentives in reinforcement learning. *Operations Research*, 72(3):983–998, 2024.
- [SW14] A. Saumard and J. A. Wellner. Log-concavity and strong log-concavity: a review. *Statistics surveys*, 8:45, 2014.
- [TI24] T. Tsuchiya and S. Ito. Fast rates in stochastic online convex optimization by exploiting the curvature of feasible sets. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.
- [VH14] R. Van Handel. Probability in high dimension. *Lecture Notes (Princeton University)*, 2(3):2–3, 2014.
- [WH18] S. Wang and L. Huang. Multi-armed bandits with compensation. *Advances in Neural Information Processing Systems*, 31, 2018.
- [WXL⁺23] H. Wang, H. Xu, C. Li, Z. Liu, and H. Wang. Incentivizing exploration in linear contextual bandits under information gap. In *Proceedings of the 17th ACM Conference on Recommender Systems*, pages 415–425, 2023.
- [ZL19] J. Zimmert and T. Lattimore. Connections between mirror descent, Thompson sampling and the information ratio. *Advances in neural information processing systems*, 32, 2019.

A Technical Lemmas

In this appendix, we introduce a series of technical lemmas that form the basis of our results. Our proofs rely on carefully analyzing how the conditional expectation of ℓ^* changes when conditioning on different signals ψ . Lemmas A.1-A.4 are the main tools we use to analyze the behavior of these conditional expectations.

Lemma A.1 says that for random variable X , if we have some signal R that is equal to X plus noise with some small probability and is just pure noise otherwise, then the conditional expectation of X given the sign of R has magnitude that is $\Omega(\epsilon)$. In the “initial exploration” phase of our algorithm we explore a new (not previously explored) direction with very small probability. Lemma A.1 implies that this exploration will lead to the conditional expectation of ℓ^* in the newly-explored direction having magnitude proportional to the probability of exploration.

Lemma A.1. *Suppose X is a real-valued K -sub-gaussian random variable such that $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = \sigma_X^2$. Let $R \sim X + N(0, 1)$ with probability ϵ and $R \sim N(0, 1)$ with probability $1 - \epsilon$. Then there exists $c_{L.A.1}$ independent of X such that*

$$|\mathbb{E}[X \mid 1_{R>0}]| \geq c_{L.A.1} \sigma_X^5 \epsilon.$$

The proof of Lemma A.1 can be found in Appendix H.1.

Lemmas A.2 and A.3 are the main technical tools that allow for us to exponentially grow the amount of exploration in any new direction. Informally, Lemma A.2 says that even if X forms only an ϵ fraction of the signal r , conditioning on the sign of r will increase the conditional expectation of X by a multiplicative factor. This lemma will be applied to the expectation of ℓ^* in the new direction we are trying to explore. Lemma A.3 says that any random variable conditioned on the sign of r cannot have conditional expectation increase by more than $O(\epsilon)$. This will be applied to the expectation of ℓ^* in all of the directions that we have already explored. These two lemmas combined allow our algorithm to multiplicatively increase the magnitude of the expectation of ℓ^* in an unexplored direction *relative* to the already explored directions.

Lemma A.2. *Let X be a K -sub-gaussian random variable satisfying $\mathbb{E}[X] = 0$ and $\mathbb{E}[X^2] = \sigma_X^2 \geq c_v$. For $Z \sim N(0, \sigma^2)$ such that $Z \perp\!\!\!\perp X$ and $\epsilon > 0$, define $r = \epsilon X + Z$. Then there exists a constant $\delta_{L.A.2}$ such that if $\epsilon/\sigma \leq \delta_{L.A.2}$,*

$$|\mathbb{E}[X \mid 1_{r>0}]| \geq \frac{\epsilon \sigma_X^2}{2\sigma\sqrt{2\pi}} \geq \frac{c_v \epsilon}{\sqrt{8\pi}\sigma} := \frac{c_{L.A.2}\epsilon}{\sigma}.$$

The proof of Lemma A.2 can be found in Appendix G.1.

Lemma A.3. *For K -sub-gaussian random variables X, Y such that $\mathbb{E}[X] = \mathbb{E}[Y] = 0$ and for $Z \sim N(0, \sigma^2)$ independent of X and Y , let $r \sim \epsilon X + Z$. Suppose $\frac{\epsilon}{\sigma} \leq \delta_{L.A.3} := \min\left(1, \frac{1}{2K\sqrt{\log(2)}}\right)$. Then there exists a constant $c_{L.A.3} > 0$ such that*

$$|\mathbb{E}[Y \mid 1_{r>0}]| \leq c_{L.A.3} \epsilon / \sigma.$$

The proof of Lemma A.3 can be found in Appendix G.2.

Lemma A.4 is a more technical lemma that allows us to better understand the distribution of ℓ^* when we condition on averages based on previous rewards. More specifically, this allows us to apply Lemma 2.2 to the random variable \mathbf{z} as described in Section 2.2. The proof of Lemma A.4 can be found in Appendix H.

Lemma A.4. *For random variable \mathbf{X} in \mathbb{R}^d , define $\mathbf{Y} = \mathbf{X} + \mathbf{W}$ where $\mathbf{W} \sim N(\mathbf{0}, \text{Diag}(\mathbf{s}))$ and \mathbf{W} is independent of \mathbf{X} . Define $\mathbf{Z}(\mathbf{Y}) = \mathbb{E}[\mathbf{X} \mid \mathbf{Y}]$. If $\min_{\|\mathbf{v}\|=1} \Pr(\langle \mathbf{X}, \mathbf{v} \rangle \geq c_d) \geq \epsilon_d$ and for all $i \in [1 : d]$, $s_i \leq c_{LA.4} := \frac{c_d^2/32}{\log(4/\epsilon_d)}$ then*

$$\min_{\|\mathbf{v}\|=1} \mathbb{E}[(\langle \mathbf{Z}(\mathbf{Y}), \mathbf{v} \rangle)^+] \geq \frac{\epsilon_d c_d}{4}.$$

The final lemma for this section says that any vector in the span of the top eigenvectors of a positive semi-definite matrix can be represented as a linear combination of these top eigenvectors with coefficients that are not too large. The proof of Lemma A.5 can be found in Appendix F.

Lemma A.5. *Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j \in \mathbb{R}^d$ such that $\|\mathbf{v}_i\| = 1$. Define $\mathbf{M} = \sum_{i=1}^j \mathbf{v}_i^{\otimes 2}$. Suppose $\mathbf{w}_1, \dots, \mathbf{w}_d$ are orthonormal eigenvectors of \mathbf{M} with corresponding eigenvalues $\lambda_1 \geq \dots \geq \lambda_d \geq 0$. Suppose $\lambda_\ell \geq \epsilon$. Then for any $\mathbf{u} \in \text{Span}(\mathbf{w}_1, \dots, \mathbf{w}_\ell)$ such that $\|\mathbf{u}\| \leq 1$, we have $\mathbf{u} = \sum_{i=1}^j c_i \mathbf{v}_i$, where $c_i := \sum_{k=1}^\ell \left(\frac{\langle \mathbf{u}, \mathbf{w}_k \rangle \langle \mathbf{v}_i, \mathbf{w}_k \rangle}{\lambda_k} \right)$. Furthermore, $\sum_{i=1}^j c_i^2 \leq \frac{1}{\epsilon}$.*

B Discussion on Assumptions

Here we illustrate why the first two conditions in Assumption 1.2 are important, by presenting an example where each is absent.

To see why Condition 1 (that ℓ^* is not confined to any half-space) is necessary, suppose $d = 2$ and that $\ell^*_1 \sim \text{Uniform}(0.5, 2)$ and $\ell^*_2 \sim \text{Uniform}(-1, 1)$ and $\ell^*_1 \perp \ell^*_2$. This example violates Condition 1, because $\Pr(\langle -\mathbf{e}_1, \ell^* \rangle \geq 0) = 0$. However, the only BIC action will for every t be $\mathbf{A}^{(t)} = \mathbf{e}_1$. This is because $\mathbb{E}[\ell^*_2] = 0$, and $\mathbb{E}[\ell^*_1 \mid \Psi] > 0$ for any signal Ψ based on any historical actions and rewards. Condition 1 is also not overly restrictive so as to make λ -

spectral exploration in $\text{poly}(d)$ steps trivial. For example, suppose $\ell^*_1 = \begin{cases} -1 & \text{w.p. } e^{-d} \\ 1 & \text{w.p. } 1 - e^{-d} \end{cases}$ and $\ell^*_2 \sim \text{Uniform}(-1, 1)$. This prior distribution satisfies Condition 1, but there is still an exponentially small probability that the optimal action is not $\mathbf{A}^* = (1, 0)$. Therefore, in this example it is not trivial to guarantee λ -spectral exploration in a polynomial number of samples.

To see why Condition 2 (that ℓ^* has non-degenerate covariance) is needed, consider the following example. Suppose $d = 2$ and that the coordinates of ℓ^* are independent with:

$$\ell^*_1 = \begin{cases} -1 & \text{w.p. } e^{-d} \\ 1 & \text{w.p. } 1 - e^{-d} \end{cases}, \quad \ell^*_2 = \begin{cases} -2 & \text{w.p. } e^{-d} \\ 0 & \text{w.p. } 1 - 2e^{-d} \\ 2 & \text{w.p. } e^{-d} \end{cases}.$$

Intuitively, this example once again requires exponentially many samples to λ -spectrally explore, because in order to explore ℓ^*_2 we need to find a signal based on the history where the expectation of ℓ^*_2 is not exponentially smaller than the expectation of ℓ^*_1 . However, we cannot do exponential growth on the conditional expectation of ℓ^*_2 , as we would need exponentially many samples to sufficiently increase the conditional expectation of ℓ^*_2 . Similarly, we would need exponentially many samples to decrease the conditional expectation of ℓ^*_1 .

We also note that any sub-gaussian random variable X satisfying $\mathbb{P}(|X| > t) \leq 2e^{-t^2/K^2}$ for all $t \geq 0$ (as in Condition 3) satisfies the following bounds on the moments of X :

$$\mathbb{E}[X] \leq \mathbb{E}[|X|] = \int_0^\infty \mathbb{P}(|X| > t) dt \leq \int_0^\infty 2e^{-t^2/K^2} dt = K\sqrt{\pi}, \quad (3)$$

$$\mathbb{E}[X^2] = \int_0^\infty \mathbb{P}(X^2 > t) dt \leq \int_0^\infty 2e^{-t/K^2} dt = 2K^2. \quad (4)$$

C Proof of Lemma 3.1

Proof of Lemma 3.1. Recall that the input parameters from Algorithm 5 come from their use in Algorithm 4. We compute as follows, with $\stackrel{d}{=}$ indicating equality in distribution.

$$\begin{aligned} \hat{y}_\ell &= \lambda c_{L,A.4} \sum_{t'=0}^{\frac{1}{\lambda c_{L,A.4}}-1} \sum_{k=1}^j \frac{\langle \mathbf{v}_k, \mathbf{w}_\ell \rangle}{\lambda_\ell} q_k^{t'} \\ &\stackrel{d}{=} \lambda c_{L,A.4} \sum_{t'=0}^{\frac{1}{\lambda c_{L,A.4}}-1} \sum_{k=1}^j \frac{\langle \mathbf{v}_k, \mathbf{w}_\ell \rangle}{\lambda_\ell} \langle \mathbf{v}_k, \ell^* \rangle + \lambda c_{L,A.4} \sum_{t'=0}^{\frac{1}{\lambda c_{L,A.4}}-1} \sum_{k=1}^j \frac{\langle \mathbf{v}_k, \mathbf{w}_\ell \rangle}{\lambda_\ell} N(0, 1) \\ &\stackrel{d}{=} \left\langle \ell^*, \left(\lambda c_{L,A.4} \sum_{t'=0}^{\frac{1}{\lambda c_{L,A.4}}-1} \sum_{k=1}^j \frac{\langle \mathbf{v}_k, \mathbf{w}_\ell \rangle}{\lambda_\ell} \mathbf{v}_k \right) \right\rangle + N \left(0, \lambda c_{L,A.4} \sum_{k=1}^j \frac{\langle \mathbf{v}_k, \mathbf{w}_\ell \rangle^2}{\lambda_\ell^2} \right) \\ &\stackrel{d}{=} \left\langle \ell^*, \left(\lambda c_{L,A.4} \sum_{t'=0}^{\frac{1}{\lambda c_{L,A.4}}-1} \mathbf{w}_\ell \right) \right\rangle + N \left(0, \frac{\lambda c_{L,A.4}}{\lambda_\ell^2} \mathbf{w}_\ell^\top \mathbf{M} \mathbf{w}_\ell \right) \quad [\text{Lemma A.5 } (\mathbf{u} = \mathbf{w}_\ell)] \\ &\stackrel{d}{=} x_\ell^* + N \left(0, c_{L,A.4} \frac{\lambda}{\lambda_\ell} \right). \end{aligned} \quad (5)$$

We will apply Lemma A.4 with $\mathbf{X} = \mathbf{x}^*$, $\mathbf{Y} = \hat{\mathbf{y}}$, and $\mathbf{Z}(\mathbf{Y}) = \mathbf{z}(\hat{\mathbf{y}})$. As shown above, (and using that $\frac{\lambda}{\lambda_\ell} \leq 1$ for all $\ell \leq \ell_\lambda$) we have that $\mathbf{Y} - \mathbf{X}$ has the appropriate distribution. The

last thing we need to show is that

$$\begin{aligned}
\min_{\|\mathbf{q}\|=1} \Pr(\langle \mathbf{X}, \mathbf{q} \rangle \geq c_d) &= \min_{\|\mathbf{q}\|=1} \Pr\left(\sum_{\ell=1}^{\ell_\lambda} \langle \boldsymbol{\ell}^*, \mathbf{w}_\ell \rangle q_\ell \geq c_d\right) \\
&= \min_{\|\mathbf{q}\|=1} \Pr\left(\left\langle \boldsymbol{\ell}^*, \left(\sum_{\ell=1}^{\ell_\lambda} q_\ell \mathbf{w}_\ell\right) \right\rangle \geq c_d\right) \\
&\geq \min_{\|\mathbf{v}\|=1} \Pr(\langle \boldsymbol{\ell}^*, \mathbf{v} \rangle \geq c_d) \\
&\geq \epsilon_d.
\end{aligned}
\tag{Assumption 1}.$$

This means we can apply Lemma A.4 to get that

$$\min_{\|\mathbf{v}\|=1} \mathbb{E}[\langle \mathbf{z}(\hat{\mathbf{y}}), \mathbf{v} \rangle^+] \geq \frac{\epsilon_d c_d}{4}.$$

We have therefore shown that $\mathbf{z}(\hat{\mathbf{y}})$ satisfies the assumptions of Lemma 2.2 with $\epsilon = \frac{\epsilon_d c_d}{4}$. \square

D Proof of Lemma 2.1

Proof of Lemma 2.1. The BIC optimal action given ψ is

$$\begin{aligned}
\mathbf{A}^* &= \arg \max_{\mathbf{A} \in S^{d-1}} \mathbb{E}[\langle \mathbf{A}, \boldsymbol{\ell}^* \rangle \mid \psi] \\
&= \arg \max_{\mathbf{A} \in S^{d-1}} \langle \mathbf{A}, \mathbb{E}[\boldsymbol{\ell}^* \mid \psi] \rangle.
\end{aligned}$$

Therefore, the BIC action is $\mathbf{A}^* = \frac{\mathbb{E}[\boldsymbol{\ell}^* \mid \psi]}{\|\mathbb{E}[\boldsymbol{\ell}^* \mid \psi]\|}$ if $\|\mathbb{E}[\boldsymbol{\ell}^* \mid \psi]\| \neq 0$. If $\|\mathbb{E}[\boldsymbol{\ell}^* \mid \psi]\| = 0$, then any action is BIC including \mathbf{v} . \square

E Proof of Lemma 2.3

We will prove the following equivalent lemma.

Lemma E.1. *In the setting of Lemma 2.3,*

$$\sum_{i=1}^{\ell} \lambda'_i \leq (1 - \epsilon/2) + \sum_{i=1}^{\ell} \lambda_i$$

We first observe that Lemma E.1 implies the desired Lemma 2.3.

Proof of Lemma 2.3. By linearity of trace,

$$\sum_{i=1}^d \lambda'_i = 1 + \sum_{i=1}^d \lambda_i.$$

Therefore, Lemma E.1 implies the desired result that

$$\sum_{i=\ell+1}^d \lambda'_i \geq \epsilon/2 + \sum_{i=\ell+1}^d \lambda_i. \quad \square$$

Proof of Lemma E.1. First, note the following, where the max is over $\mathbf{x}_1, \dots, \mathbf{x}_\ell$ that are orthonormal.

$$\sum_{i=1}^{\ell} \lambda'_i = \max_{\mathbf{x}_1, \dots, \mathbf{x}_\ell} \sum_{i=1}^{\ell} \mathbf{x}_i^T \mathbf{M}' \mathbf{x}_i = \max_{\mathbf{x}_1, \dots, \mathbf{x}_\ell} \left(\sum_{i=1}^{\ell} \mathbf{x}_i^T \mathbf{M} \mathbf{x}_i + \sum_{i=1}^{\ell} \langle \mathbf{x}_i, \mathbf{u} \rangle^2 \right). \quad (6)$$

Define

$$\mathbf{x}_1^*, \dots, \mathbf{x}_\ell^* = \arg \max_{\mathbf{x}_1, \dots, \mathbf{x}_\ell} \sum_{i=1}^{\ell} \mathbf{x}_i^T \mathbf{M}' \mathbf{x}_i. \quad (7)$$

We will now prove (by contradiction) that for all $i \leq \ell$, $\|\mathcal{P}_{S^\perp}(\mathbf{x}_i^*)\|_2^2 \leq \frac{\epsilon^2}{100d^2}$. Suppose that there exists some $i' \in 1, \dots, \ell$ such that $\|\mathcal{P}_{S^\perp}(\mathbf{x}_{i'}^*)\|_2^2 > \frac{\epsilon^2}{100d^2}$. Then we find

$$\begin{aligned} \sum_{i=1}^{\ell} (\mathbf{x}_i^*)^T \mathbf{M}' \mathbf{x}_i^* &= \sum_{i=1}^{\ell} (\mathbf{x}_i^*)^T \mathbf{M} \mathbf{x}_i^* + \sum_{i=1}^{\ell} \langle \mathbf{x}_i^*, \mathbf{u} \rangle^2 \\ &\leq 1 + \sum_{i=1}^{\ell} (\mathbf{x}_i^*)^T \mathbf{M} \mathbf{x}_i^* \quad [\mathbf{x}_i^* \text{ orthonormal so } \sum_{i=1}^{\ell} \langle \mathbf{x}_i^*, \mathbf{u} \rangle^2 \leq \|\mathbf{u}\|^2 \leq 1] \\ &= 1 + \sum_{i=1}^{\ell} (\mathcal{P}_{S^\perp}(\mathbf{x}_i^*)^T \mathbf{M} \mathcal{P}_{S^\perp}(\mathbf{x}_i^*) + \mathcal{P}_S(\mathbf{x}_i^*)^T \mathbf{M} \mathcal{P}_S(\mathbf{x}_i^*)) \\ &= 1 + \sum_{i=1}^{\ell} \mathcal{P}_{S^\perp}(\mathbf{x}_i^*)^T \mathbf{M} \mathcal{P}_{S^\perp}(\mathbf{x}_i^*) + \sum_{i=1}^{\ell} \mathcal{P}_S(\mathbf{x}_i^*)^T \mathbf{M} \mathcal{P}_S(\mathbf{x}_i^*) \\ &= 1 + d\epsilon + \sum_{i=1}^{\ell} \mathcal{P}_S(\mathbf{x}_i^*)^T \mathbf{M} \mathcal{P}_S(\mathbf{x}_i^*) \quad [S^\perp = \text{span of vectors with values } \leq \epsilon] \\ &= 1 + d\epsilon + \sum_{i=1}^{\ell} \left(\sum_{k=1}^{\ell_\epsilon} \langle \mathbf{x}_i^*, \mathbf{w}_k \rangle \mathbf{w}_k \right)^T \mathbf{M} \left(\sum_{k=1}^{\ell_\epsilon} \langle \mathbf{x}_i^*, \mathbf{w}_k \rangle \mathbf{w}_k \right) \\ &= 1 + d\epsilon + \sum_{i=1}^{\ell} \sum_{k=1}^{\ell_\epsilon} \langle \mathbf{x}_i^*, \mathbf{w}_k \rangle^2 \mathbf{w}_k^T \mathbf{M} \mathbf{w}_k \\ &= 1 + d\epsilon + \sum_{i=1}^{\ell} \sum_{k=1}^{\ell_\epsilon} \langle \mathbf{x}_i^*, \mathbf{w}_k \rangle^2 \lambda_k. \end{aligned} \quad (8)$$

Because \mathbf{x}_i^* are orthonormal, we know that $\sum_{i=1}^{\ell} \langle \mathbf{x}_i^*, \mathbf{w}_k \rangle^2 \leq \|\mathbf{w}_k\|_2^2 = 1$. Because we assumed that $\|\mathcal{P}_{S^\perp}(\mathbf{x}_{i'}^*)\|_2^2 > \frac{\epsilon^2}{100d^2}$, we know that $\sum_{i=1}^{\ell} \sum_{k=1}^{\ell_\epsilon} \langle \mathbf{x}_i^*, \mathbf{w}_k \rangle^2 = \sum_{i=1}^{\ell} \|\mathcal{P}_S(\mathbf{x}_i^*)\|^2 \leq \ell - \frac{\epsilon^2}{100d^2}$. Combining these two statements with the fact that λ_k is decreasing in k ,

$$\sum_{i=1}^{\ell} \sum_{k=1}^{\ell_\epsilon} \langle \mathbf{x}_i^*, \mathbf{w}_k \rangle^2 \lambda_k = \sum_{k=1}^{\ell_\epsilon} \sum_{i=1}^{\ell} \langle \mathbf{x}_i^*, \mathbf{w}_k \rangle^2 \lambda_k \leq \left(1 - \frac{\epsilon^2}{100d^2} \right) \lambda_\ell + \sum_{i=1}^{\ell-1} \lambda_i.$$

Continuing where we left off with Equation (8), we have that

$$\begin{aligned}
&= 1 + d\epsilon + \sum_{i=1}^{\ell} \sum_{k=1}^{\ell_{\epsilon}} \langle \mathbf{x}_i^*, \mathbf{w}_k \rangle^2 \lambda_k \leq 1 + d\epsilon + \left(1 - \frac{\epsilon^2}{100d^2}\right) \lambda_{\ell} + \sum_{i=1}^{\ell-1} \lambda_i \\
&\leq 1 + d\epsilon - \frac{\epsilon^2}{100d^2} \lambda_{\ell} + \sum_{i=1}^{\ell} \lambda_i \\
&\leq 1 + d\epsilon - 2d + \sum_{i=1}^{\ell} \lambda_i && [\lambda_{\ell} \geq \frac{200d^3}{\epsilon^2}] \\
&< \sum_{i=1}^{\ell} \lambda_i && [\epsilon < 1] \\
&= \sum_{i=1}^{\ell} \mathbf{w}_i^{\top} \mathbf{M} \mathbf{w}_i \leq \sum_{i=1}^{\ell} \mathbf{w}_i^{\top} \mathbf{M}' \mathbf{w}_i.
\end{aligned}$$

Therefore, we have a contradiction, as $\mathbf{x}_1^*, \dots, \mathbf{x}_{\ell}^*$ cannot be a solution to Equation (7) because these vectors are strictly beaten by $\mathbf{w}_1, \dots, \mathbf{w}_{\ell}$.

Therefore, we have shown that $\|\mathcal{P}_{S^{\perp}}(\mathbf{x}_i^*)\|_2^2 \leq \frac{\epsilon^2}{100d^2}$ for all i .

In the following equation, we define P_S as the projection matrix for projecting a vector onto

S. Now, we have that

$$\begin{aligned}
\sum_{i=1}^{\ell} \lambda'_i &= \sum_{i=1}^{\ell} (\mathbf{x}_i^*)^\top \mathbf{M} \mathbf{x}_i^* + \sum_{i=1}^{\ell} \langle \mathbf{x}_i^*, \mathbf{u} \rangle^2 \\
&\leq \sum_{i=1}^{\ell} \lambda_i + \sum_{i=1}^{\ell} \langle \mathbf{x}_i^*, \mathbf{u} \rangle^2 \\
&\leq \sum_{i=1}^{\ell} \lambda_i + \sum_{i=1}^{\ell} \left(\langle \mathcal{P}_S(\mathbf{x}_i^*), \mathcal{P}_S(\mathbf{u}) \rangle + \langle \mathcal{P}_{S^\perp}(\mathbf{x}_i^*), \mathcal{P}_{S^\perp}(\mathbf{u}) \rangle \right)^2 \\
&\leq \sum_{i=1}^{\ell} \lambda_i + \sum_{i=1}^{\ell} \left(|\langle \mathcal{P}_S(\mathbf{x}_i^*), \mathcal{P}_S(\mathbf{u}) \rangle| + \frac{\epsilon}{10d} \right)^2 \quad [\|\mathcal{P}_{S^\perp}(\mathbf{x}_i^*)\|_2^2 \leq \frac{\epsilon^2}{100d^2}] \\
&= \sum_{i=1}^{\ell} \lambda_i + \sum_{i=1}^{\ell} \left(\langle \mathcal{P}_S(\mathbf{x}_i^*), \mathcal{P}_S(\mathbf{u}) \rangle^2 + \frac{\epsilon}{5d} |\langle \mathcal{P}_S(\mathbf{x}_i^*), \mathcal{P}_S(\mathbf{u}) \rangle| + \frac{\epsilon^2}{100d^2} \right) \\
&\leq \sum_{i=1}^{\ell} \lambda_i + \sum_{i=1}^{\ell} \left(\langle \mathcal{P}_S(\mathbf{x}_i^*), \mathcal{P}_S(\mathbf{u}) \rangle^2 + \frac{\epsilon}{5d} + \frac{\epsilon^2}{100d^2} \right) \\
&\leq \sum_{i=1}^{\ell} \lambda_i + \frac{\epsilon}{5} + \frac{\epsilon^2}{100d} + \sum_{i=1}^{\ell} \langle \mathcal{P}_S(\mathbf{x}_i^*), \mathcal{P}_S(\mathbf{u}) \rangle^2 \\
&= \sum_{i=1}^{\ell} \lambda_i + \frac{\epsilon}{5} + \frac{\epsilon^2}{100d} + \sum_{i=1}^{\ell} ((\mathbf{x}_i^*)^\top P_S^\top P_S \mathbf{u})^2 \\
&= \sum_{i=1}^{\ell} \lambda_i + \frac{\epsilon}{5} + \frac{\epsilon^2}{100d} + \sum_{i=1}^{\ell} ((\mathbf{x}_i^*)^\top P_S \mathbf{u})^2 \\
&= \sum_{i=1}^{\ell} \lambda_i + \frac{\epsilon}{5} + \frac{\epsilon^2}{100d} + \sum_{i=1}^{\ell} \langle \mathbf{x}_i^*, \mathcal{P}_S(\mathbf{u}) \rangle^2 \\
&\leq \sum_{i=1}^{\ell} \lambda_i + \frac{\epsilon}{5} + \frac{\epsilon^2}{100d} + \|\mathcal{P}_S(\mathbf{u})\|_2^2 \\
&\leq \sum_{i=1}^{\ell} \lambda_i + \frac{\epsilon}{5} + \frac{\epsilon^2}{100d} + 1 - \epsilon \\
&\leq \sum_{i=1}^{\ell} \lambda_i + (1 - \epsilon/2).
\end{aligned}$$

This completes the proof of Lemma E.1. □

F Proof of Lemma A.5

Proof of Lemma A.5. Define $\mathbf{A} \in \mathbb{R}^{j \times d}$ with rows corresponding to $\mathbf{v}_1, \dots, \mathbf{v}_j$. Then $\mathbf{M} = \mathbf{A}^\top \mathbf{A}$. We want to write $\mathbf{u} = \mathbf{A}^\top \mathbf{c}$ for $\mathbf{c} \in \mathbb{R}^j$.

Because $\mathbf{w}_1, \dots, \mathbf{w}_\ell$ are orthonormal and $\mathbf{u} \in \text{Span}(\mathbf{w}_1, \dots, \mathbf{w}_\ell)$, we have that

$$\mathbf{u} = \sum_{i=1}^{\ell} \langle \mathbf{u}, \mathbf{w}_i \rangle \mathbf{w}_i$$

Because \mathbf{w}_i is an eigenvector of \mathbf{M} with eigenvalue λ_i , we know that for any $i \leq \ell$

$$\lambda_i \mathbf{w}_i = \mathbf{M} \mathbf{w}_i = \mathbf{A}^\top \mathbf{A} \mathbf{w}_i$$

Rearranging terms and multiplying both sides by $\langle \mathbf{u}, \mathbf{w}_i \rangle$, we have that for any $i \leq \ell$

$$\langle \mathbf{u}, \mathbf{w}_i \rangle \mathbf{w}_i = \mathbf{A}^\top \left(\frac{\langle \mathbf{u}, \mathbf{w}_i \rangle \mathbf{A} \mathbf{w}_i}{\lambda_i} \right).$$

Therefore, we have that

$$\mathbf{u} = \sum_{i=1}^{\ell} \langle \mathbf{u}, \mathbf{w}_i \rangle \mathbf{w}_i = \sum_{i=1}^{\ell} \mathbf{A}^\top \left(\frac{\langle \mathbf{u}, \mathbf{w}_i \rangle \mathbf{A} \mathbf{w}_i}{\lambda_i} \right) = \mathbf{A}^\top \sum_{i=1}^{\ell} \left(\frac{\langle \mathbf{u}, \mathbf{w}_i \rangle \mathbf{A} \mathbf{w}_i}{\lambda_i} \right).$$

Now, we can define

$$\mathbf{c} = \sum_{i=1}^{\ell} \left(\frac{\langle \mathbf{u}, \mathbf{w}_i \rangle \mathbf{A} \mathbf{w}_i}{\lambda_i} \right) = \mathbf{A} \sum_{i=1}^{\ell} \frac{\langle \mathbf{u}, \mathbf{w}_i \rangle \mathbf{w}_i}{\lambda_i}.$$

This implies that

$$\begin{aligned} \|\mathbf{c}\|_2^2 &= \left\| \mathbf{A} \sum_{i=1}^{\ell} \frac{\langle \mathbf{u}, \mathbf{w}_i \rangle \mathbf{w}_i}{\lambda_i} \right\|_2^2 \\ &= \left(\sum_{i=1}^{\ell} \frac{\langle \mathbf{u}, \mathbf{w}_i \rangle \mathbf{w}_i}{\lambda_i} \right)^\top \mathbf{M} \left(\sum_{i=1}^{\ell} \frac{\langle \mathbf{u}, \mathbf{w}_i \rangle \mathbf{w}_i}{\lambda_i} \right) \\ &= \sum_{i=1}^{\ell} \lambda_i \left(\frac{\langle \mathbf{u}, \mathbf{w}_i \rangle}{\lambda_i} \right)^2 && [\mathbf{w}_i^\top \mathbf{M} \mathbf{w}_i = \lambda_i] \\ &= \sum_{i=1}^{\ell} \frac{\langle \mathbf{u}, \mathbf{w}_i \rangle^2}{\lambda_i} \\ &\leq \frac{\|\mathbf{u}\|_2^2}{\epsilon} \\ &\leq \frac{1}{\epsilon}. \end{aligned}$$

This vector \mathbf{c} therefore satisfies the desired properties. □

G Proof of Lemma 2.2

We begin by proving the following lemma.

Lemma G.1. *Let μ be a probability distribution on \mathbb{R}^d with finite first moment and suppose*

$$\min_{\|\mathbf{v}\|=1} \mathbb{E}^{\mathbf{x} \sim \mu}[\langle \mathbf{v}, \mathbf{x} \rangle_+] \geq \epsilon.$$

Then for any \mathbf{w} with $\|\mathbf{w}\| < \epsilon$ there is a $[0, 1]$ -valued measurable function f such that $\mathbb{E}[\mathbf{x}f(\mathbf{x})] = \mathbf{w}$.

Proof of Lemma G.1. Let K be the set of possible vectors $\mathbb{E}[\mathbf{x}f(\mathbf{x})]$ where f is a $[0, 1]$ -valued measurable function. Then for any $\mathbf{a}, \mathbf{b} \in K$, there exist corresponding $[0, 1]$ -valued functions f^a and f^b such that $\mathbb{E}[\mathbf{x}f^a(\mathbf{x})] = \mathbf{a}$ and $\mathbb{E}[\mathbf{x}f^b(\mathbf{x})] = \mathbf{b}$. Therefore for any $t \in [0, 1]$, $t\mathbf{a} + (1-t)\mathbf{b} \in K$ because $\mathbb{E}[\mathbf{x}f^t(\mathbf{x})] = t\mathbf{a} + (1-t)\mathbf{b}$ for $f^t(\mathbf{x}) = tf^a(\mathbf{x}) + (1-t)f^b(\mathbf{x})$. This implies that K is convex.

We will now prove the desired result by contradiction. Suppose $\mathbf{w} \notin K$. Because K is convex, if $\mathbf{w} \notin K$ then there is a “separating hyperplane” unit vector \mathbf{v} such that

$$\sup_{\mathbf{u} \in K} \langle \mathbf{v}, \mathbf{u} \rangle \leq \langle \mathbf{v}, \mathbf{w} \rangle.$$

(Note that we do not argue here that K is closed.) Because by assumption we have that $\|\mathbf{w}\| < \epsilon$, this implies that

$$\sup_{\mathbf{u} \in K} \langle \mathbf{v}, \mathbf{u} \rangle < \epsilon.$$

For any \mathbf{v} , by definition of K and linearity of expectation we have that

$$\begin{aligned} \sup_{\mathbf{u} \in K} \langle \mathbf{v}, \mathbf{u} \rangle &= \sup_{f: \mathbb{R}^d \rightarrow [0,1]} \langle \mathbf{v}, \mathbb{E}^{\mathbf{x} \sim \mu}[\mathbf{x}f(\mathbf{x})] \rangle \\ &= \sup_{f: \mathbb{R}^d \rightarrow [0,1]} \mathbb{E}^{\mathbf{x} \sim \mu}[f(\mathbf{x})\langle \mathbf{v}, \mathbf{x} \rangle] \\ &= \mathbb{E}^{\mathbf{x} \sim \mu}[\langle \mathbf{v}, \mathbf{x} \rangle_+] \\ &\geq \epsilon. \end{aligned}$$

where the last line followed from the assumption of the lemma. This gives a contradiction, and therefore $\mathbf{w} \in K$ must be true. \square

Proof of Lemma 2.2. Applying the above lemma with $\mathbf{w} = \mathbf{0}$, there exists $f^0 : \mathbb{R}^d \rightarrow [0, 1]$ such that $\mathbb{E}[\mathbf{x}f^0(\mathbf{x})] = \mathbf{0}$. Define the function f' as $f'(\mathbf{x}) = \frac{f^0(\mathbf{x}) + 2\epsilon}{4 \max(\|\mathbb{E}[\mathbf{x}]\|, 1)}$. By this construction,

$$\|\mathbb{E}[\mathbf{x}f'(\mathbf{x})]\| = \left\| \frac{\mathbb{E}[\mathbf{x}f^0(\mathbf{x})] + 2\epsilon\mathbb{E}[\mathbf{x}]}{4 \max(\|\mathbb{E}[\mathbf{x}]\|, 1)} \right\| = \frac{2\epsilon \|\mathbb{E}[\mathbf{x}]\|}{4 \max(\|\mathbb{E}[\mathbf{x}]\|, 1)} \leq \frac{\epsilon}{2}.$$

Therefore, $\mathbf{w} := -\mathbb{E}[\mathbf{x}f'(\mathbf{x})]$ satisfies $\|\mathbf{w}\| < \epsilon$. Applying the above lemma again, there exists $f^w : \mathbb{R}^d \rightarrow [0, 1]$ such that $\mathbb{E}[\mathbf{x}f^w(\mathbf{x})] = -\mathbb{E}[\mathbf{x}f'(\mathbf{x})]$. Now define $f(\mathbf{x}) = \frac{f'(\mathbf{x}) + f^w(\mathbf{x})}{2}$.

By construction, we have that $1 \geq f(\mathbf{x}) \geq \frac{f'(\mathbf{x})}{2} \geq \frac{\epsilon}{4 \max(\|\mathbb{E}[\mathbf{x}]\|, 1)}$. Furthermore, by linearity of expectation and the construction of f^w we have that

$$\mathbb{E}[\mathbf{x}f(\mathbf{x})] = \mathbb{E}\left[\mathbf{x} \frac{f'(\mathbf{x}) + f^w(\mathbf{x})}{2}\right] = \frac{1}{2} (\mathbb{E}[\mathbf{x}f'(\mathbf{x})] + \mathbb{E}[\mathbf{x}f^w(\mathbf{x})]) = \frac{1}{2} (\mathbb{E}[\mathbf{x}f'(\mathbf{x})] - \mathbb{E}[\mathbf{x}f'(\mathbf{x})]) = 0.$$

Therefore, $f(\mathbf{x})$ is a $\left[\frac{\epsilon}{4 \max(\|\mathbb{E}[\mathbf{x}]\|, 1)}, 1\right]$ -valued function that satisfies $\mathbb{E}[\mathbf{x}f(\mathbf{x})] = 0$ as desired. \square

G.1 Proof of Lemma A.2

Lemma G.2. *Let $\Phi^C(x) = \mathbb{P}(Z > x)$ for $Z \sim N(0, 1)$. Then for $|x| \leq 1$,*

$$\left| \Phi^C(x) - \left(\frac{1}{2} - \frac{1}{\sqrt{2\pi}}x \right) \right| \leq |x^3|/15. \quad (9)$$

Proof of Lemma G.2. A third order Taylor expansion shows the error is at most

$$|x^3| \cdot \frac{\sup_{|y| \leq 1} |(\Phi^C)'''(y)|}{6}.$$

For $|y| \leq 1$ we easily compute

$$|(\Phi^C)'''(y)| = \frac{|y^2 - 1| e^{-y^2/2}}{\sqrt{2\pi}} \leq 1/\sqrt{2\pi} \leq 2/5. \quad \square$$

Lemma G.3. *If X is K -sub-gaussian, then for any event E and any $\mathbb{P}(E) \geq a > 0$, we have*

$$\mathbb{E}[X^2 1_E] \leq K^2 \Pr(E) \log(2/a) + K^2 a, \quad (10)$$

$$\mathbb{E}[X 1_E] \leq O(\Pr(E) \log(1/a)). \quad (11)$$

Proof of Lemma G.3. We prove both estimates using the tail-sum formula. For the truncated second moment,

$$\begin{aligned} \mathbb{E}[X^2 1_E] &= \int_0^\infty \Pr(|X^2 1_E| \geq t) dt \\ &\leq \int_0^\infty \min(\Pr(E), \Pr(X^2 \geq t)) dt \\ &\leq \int_0^\infty \min(\Pr(E), 2e^{-t/K^2}) dt \\ &\leq \Pr(E) \log(2/a) K^2 + \int_{\log(2/a) K^2}^\infty 2e^{-t/K^2} dt \\ &= K^2 \Pr(E) \log(2/a) + K^2 a. \end{aligned}$$

Similarly for the truncated first moment,

$$\begin{aligned}
\mathbb{E}[X1_E] &= \int_0^\infty \Pr(|X1_E| \geq t) dt \\
&\leq \int_0^\infty \min(\Pr(E), \Pr(X \geq t)) dt \\
&\leq \int_0^\infty \min(\Pr(E), 2e^{-t^2/K^2}) dt \\
&\leq \Pr(E) \sqrt{\log(3/a)} K + \int_{\sqrt{\log(3/a)} K}^\infty 2e^{-t^2/K^2} dt \\
&= O(\Pr(E) \log(1/a)). \quad \square
\end{aligned}$$

Proof of Lemma A.2. Let $d\mu_X$ be the law of X and $d\mu_{X|r>0}$ be the conditional law of X given the event $r > 0$. Then by Bayes rule,

$$\begin{aligned}
d\mu_{X|r>0}(x) &= \frac{\Pr(r > 0 \mid X = x) d\mu_X(x)}{\Pr(r > 0)} \\
&= \frac{\Pr(N(\epsilon x, \sigma^2) > 0) d\mu_X(x)}{\Pr(r > 0)} \\
&= \frac{\Phi^C(-\epsilon x/\sigma) d\mu_X(x)}{\Pr(r > 0)}. \tag{12}
\end{aligned}$$

Note that

$$\mathbb{E}[X \mid r > 0] = \int_{-\infty}^\infty x d\mu_{X|r>0}(x) = \frac{1}{\Pr(r > 0)} \int_{-\infty}^\infty x \Phi^C(-\epsilon x/\sigma) d\mu_X(x).$$

Since $\Pr(r > 0) \leq 1$ it suffices to lower-bound the latter integral by a suitable positive value. As long as $\epsilon/\sigma \leq \delta_{L\text{A.2}} \leq \sqrt{\frac{1}{4\sigma_X^2 \sqrt{2\pi}}}$, we have

$$\begin{aligned}
&\int_{-\infty}^\infty x \Phi^C(-\epsilon x/\sigma) d\mu_X(x) \\
&= \int_{-\infty}^\infty x \left(\frac{1}{2} + \frac{1}{\sqrt{2\pi}} \frac{\epsilon x}{\sigma} \right) d\mu_X(x) \\
&\quad + \int_{-\infty}^\infty x \left(\Phi^C\left(\frac{-\epsilon x}{\sigma}\right) - \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \frac{\epsilon x}{\sigma} \right) d\mu_X(x) \\
&= \left(\frac{\epsilon \mathbb{E}[X^2]}{\sigma \sqrt{2\pi}} + \int_{-\infty}^\infty x \left(\Phi^C\left(\frac{-\epsilon x}{\sigma}\right) - \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \frac{\epsilon x}{\sigma} \right) d\mu_X(x) \right) \quad [\mathbb{E}[X] = 0] \\
&\geq \left(\frac{\epsilon \sigma_X^2}{\sigma \sqrt{2\pi}} - \frac{2 \mathbb{E}[X^4] \epsilon^3}{\sigma^3} \right) \quad [\text{Inequality (13) Below}] \\
&\geq \frac{\epsilon \sigma_X^2}{2\sigma \sqrt{2\pi}}. \quad \left[\frac{\epsilon^2}{\sigma^2} \leq \frac{\sigma_X^2}{4 \mathbb{E}[X^4] \sqrt{2\pi}} \right]
\end{aligned}$$

Above we used the following estimate (13). For sufficiently small $\delta_{L\text{A.2}}$,

$$\begin{aligned}
& \int_{-\infty}^{\infty} x \left(\Phi^C\left(\frac{-\epsilon x}{\sigma}\right) - \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \frac{\epsilon x}{\sigma} \right) d\mu_X(x) \\
&= \int_{|x| \leq \frac{\sigma}{10\epsilon}} x \left(\Phi^C\left(\frac{-\epsilon x}{\sigma}\right) - \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \frac{\epsilon x}{\sigma} \right) d\mu_X(x) + \int_{|x| > \frac{\sigma}{10\epsilon}} x \left(\Phi^C\left(\frac{-\epsilon x}{\sigma}\right) - \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \frac{\epsilon x}{\sigma} \right) d\mu_X(x) \\
&\geq \int_{|x| \leq \frac{\sigma}{10\epsilon}} x \left(\Phi^C\left(\frac{-\epsilon x}{\sigma}\right) - \frac{1}{2} - \frac{1}{\sqrt{2\pi}} \frac{\epsilon x}{\sigma} \right) d\mu_X(x) - \int_{|x| > \frac{\sigma}{10\epsilon}} \left(\frac{|x|}{2} + \frac{\epsilon x^2}{\sigma\sqrt{2\pi}} \right) d\mu_X(x) \\
&\geq - \int_{|x| \leq \frac{\sigma}{10\epsilon}} |x| \left| \frac{\epsilon x}{\sigma} \right|^3 d\mu_X(x) - \int_{|x| > \frac{\sigma}{10\epsilon}} \left(\frac{|x|}{2} + \frac{\epsilon x^2}{\sigma\sqrt{2\pi}} \right) d\mu_X(x) \quad [\text{Lemma G.2}] \\
&= - \frac{\mathbb{E}[X^4]\epsilon^3}{\sigma^3} - \int_{|x| > \frac{\sigma}{10\epsilon}} \left(\frac{|x|}{2} + \frac{\epsilon x^2}{\sigma\sqrt{2\pi}} \right) d\mu_X(x) \\
&\geq - \frac{\mathbb{E}[X^4]\epsilon^3}{\sigma^3} - \int_{|x| > \frac{\sigma}{10\epsilon}} x^2 d\mu_X(x) \quad [\text{if } \epsilon/\sigma \leq \delta_{L\text{A.2}} \leq 1/10] \\
&\geq - \frac{\mathbb{E}[X^4]\epsilon^3}{\sigma^3} - (K^2 \Pr(E) \log(2/a) + K^2 a) \quad [\text{Lemma G.3, } E := \{|x| > \frac{\sigma}{10\epsilon}\}, a = 2e^{-\frac{\sigma^2}{100\epsilon^2 K^2}}] \\
&\geq - \frac{\mathbb{E}[X^4]\epsilon^3}{\sigma^3} - 2K^2 e^{-\frac{\sigma^2}{100K^2\epsilon^2}} \left(\frac{\sigma^2}{100\epsilon^2 K^2} + 1 \right) \quad [\Pr(E) \leq 2e^{-\sigma^2/(100\epsilon^2 K^2)}] \\
&\geq - \frac{2\mathbb{E}[X^4]\epsilon^3}{\sigma^3} \quad [2K^2 e^{-\frac{\sigma^2}{100K^2\epsilon^2}} \left(\frac{\sigma^2}{100\epsilon^2 K^2} + 1 \right) \leq \frac{(\sigma_X^2)^2 \epsilon^3}{\sigma^3} \leq \frac{\mathbb{E}[X^4]\epsilon^3}{\sigma^3}] \\
&\quad (13)
\end{aligned}$$

where the second to last line holds for sufficiently small ϵ/σ . \square

G.2 Proof of Lemma A.3

Proof of Lemma A.3. First, we observe that $\mathbb{E}[Y \mid r > 0] = \frac{\mathbb{E}[Y \mathbf{1}\{r > 0\}]}{\mathbb{P}(r > 0)}$. If $\frac{\epsilon}{\sigma} K \sqrt{\log(4)} \leq 1/2$, we also have that

$$\begin{aligned}
& \Pr(r > 0) \\
&\geq \Pr(r > 0 \mid X \geq -K\sqrt{\log(4)}) \Pr(X \geq -K\sqrt{\log(4)}) \\
&\geq \Phi^C\left(\frac{\epsilon}{\sigma} K \sqrt{\log(4)}\right) \Pr(X \geq -K\sqrt{\log(4)}) \\
&\geq \Phi^C\left(\frac{\epsilon}{\sigma} K \sqrt{\log(4)}\right) \left(1 - 2e^{-K^2 \log(4)/K^2}\right) \\
&\geq \Phi^C(1/2) 1/2 \quad \left[\frac{\epsilon}{\sigma} K \sqrt{\log(4)} \leq 1/2\right] \\
&\geq 1/8.
\end{aligned}$$

Therefore, it is sufficient to upper bound $|\mathbb{E}[Y \mathbf{1}\{r > 0\}]|$. By law of total expectation,

$$\mathbb{E}[Y \mathbf{1}\{r > 0\}] = \mathbb{E}[\mathbb{E}[Y \mathbf{1}\{r > 0\} \mid X]] = \mathbb{E}\left[\mathbb{E}[Y \mid X] \mathbb{P}(Z > -\frac{\epsilon}{\sigma} X)\right].$$

Define $Q(X) = \mathbb{E}[Y | X]$. Because Y is sub-gaussian, the random variable $Q(X)$ must also be sub-gaussian with parameter $\sqrt{18}K$ (see [VH14, Exercise 3.1]). Furthermore, $\mathbb{E}[Q(X)] = \mathbb{E}[\mathbb{E}[Y | X]] = \mathbb{E}[Y] = 0$. Now, we have the following

$$\begin{aligned}
& |\mathbb{E}[Y \mathbf{1}\{r > 0\}]| \\
&= \left| \mathbb{E} \left[\mathbb{E}[Y | X] \mathbb{P}(Z > -\frac{\epsilon}{\sigma} X) \right] \right| \\
&= \left| \int_{-\infty}^{\infty} Q(x) \Phi^C(-\frac{\epsilon x}{\sigma}) d\mu_X(x) \right| \\
&= \left| \int_{-\infty}^{\infty} \frac{1}{2} Q(x) d\mu_X(x) + \int_{-\infty}^{\infty} Q(x) \left(\Phi^C(-\frac{\epsilon x}{\sigma}) - \frac{1}{2} \right) d\mu_X(x) \right| \\
&= \left| \frac{1}{2} \mathbb{E}[Q(X)] + \int_{-\infty}^{\infty} Q(x) \left(\Phi^C(-\frac{\epsilon x}{\sigma}) - \frac{1}{2} \right) d\mu_X(x) \right| \\
&= \left| \int_{-\infty}^{\infty} Q(x) \left(\Phi^C(-\frac{\epsilon x}{\sigma}) - \frac{1}{2} \right) d\mu_X(x) \right| \\
&\leq \int_{|x| \leq \frac{\sigma}{10\epsilon}} |Q(x)| \left(\frac{|\epsilon x|}{\sigma \sqrt{2\pi}} + \left| \frac{\epsilon x}{\sigma} \right|^3 \right) d\mu_X(x) \\
&\quad + \int_{|x| > \frac{\sigma}{10\epsilon}} Q(x) \left(\Phi^C(-\frac{\epsilon x}{\sigma}) - \frac{1}{2} \right) d\mu_X(x) \quad [\text{Lemma G.2}] \\
&\leq \frac{\epsilon}{\sigma} \int_{-\infty}^{\infty} |Q(x)| \left(\frac{|x|}{\sqrt{2\pi}} + |x^3| \right) d\mu_X(x) + \frac{1}{2} \int_{|x| > \frac{\sigma}{10\epsilon}} |Q(x)| d\mu_X(x) \quad [\epsilon/\sigma \leq 1] \\
&\leq \frac{\epsilon}{\sigma} \sqrt{\mathbb{E}[Q(X)^2] \mathbb{E} \left[\left(\frac{|x|}{\sqrt{2\pi}} + |x^3| \right)^2 \right]} + \frac{1}{2} O \left(\frac{\sigma}{10\epsilon K} (2e^{-\frac{\sigma^2}{100\epsilon^2 K^2}}) \right) \quad [\text{Lemma G.3 Equation (11)}] \\
&\leq O(\epsilon/\sigma).
\end{aligned}$$

where in the second to last line we used that $\Pr(|X| > \frac{\sigma}{10\epsilon}) \leq 2e^{-\frac{\sigma^2}{100\epsilon^2 K^2}}$ and that $Q(X)$ is $K\sqrt{18}$ -sub-gaussian. The last line again uses that both X and $Q(X)$ are sub-gaussian.

Therefore, we have shown that

$$|\mathbb{E}[Y | r > 0]| = \left| \frac{\mathbb{E}[Y \mathbf{1}\{r > 0\}]}{\mathbb{P}(r > 0)} \right| \leq 8 |\mathbb{E}[Y \mathbf{1}\{r > 0\}]| = O(\epsilon/\sigma).$$

By symmetric arguments to the ones above, we have the same bound on $|\mathbb{E}[Y | r \leq 0]|$.

□

H Proof of Lemma A.4

Proof of Lemma A.4. Fix any \mathbf{v} such that $\|\mathbf{v}\| = 1$. First, we will show that

$$\Pr(|\mathbf{v} \cdot (\mathbf{Z}(\mathbf{Y}) - \mathbf{X})| \geq c_d/2) \leq \epsilon_d/2.$$

To do this, we will show that $\langle \mathbf{v}, (\mathbf{Z}(\mathbf{Y}) - \mathbf{X}) \rangle$ is a sub-gaussian random variable. Let $\hat{\mathbf{X}}$ be a draw from the distribution of $\mathbf{X} \mid \mathbf{Y}$. Then we have that

$$\langle \mathbf{v}, \hat{\mathbf{X}} - \mathbf{X} \rangle = \langle \mathbf{v}, \hat{\mathbf{X}} - \mathbf{Y} \rangle + \langle \mathbf{v}, \mathbf{Y} - \mathbf{X} \rangle.$$

By standard properties of posterior samples, $\langle \mathbf{v}, \hat{\mathbf{X}} - \mathbf{Y} \rangle$ and $\langle \mathbf{v}, \mathbf{Y} - \mathbf{X} \rangle$ are identically distributed with distribution $N(0, \sigma^2)$ for $\sigma^2 = \sum_{i=1}^d \mathbf{v}_i^2 s_i$ (here one averages over all randomness). Therefore, we have that

$$\begin{aligned} & \mathbb{E} [\exp (t \langle \mathbf{v}, \mathbf{Z}(\mathbf{Y}) - \mathbf{X} \rangle)] \\ &= \mathbb{E} \left[\exp \left(t \langle \mathbf{v}, \mathbb{E}[\hat{\mathbf{X}}] - \mathbf{X} \rangle \right) \right] \\ &\leq \mathbb{E} \left[\exp \left(t \langle \mathbf{v}, \hat{\mathbf{X}} - \mathbf{X} \rangle \right) \right] \quad [\text{Jensen}] \\ &= \mathbb{E} \left[\exp \left(t \langle \mathbf{v}, \hat{\mathbf{X}} - \mathbf{Y} + \mathbf{Y} - \mathbf{X} \rangle \right) \right] \\ &\leq \sqrt{\mathbb{E} \left[\exp \left(2t \langle \mathbf{v}, \hat{\mathbf{X}} - \mathbf{Y} \rangle \right) \right] \mathbb{E} [\exp (2t \langle \mathbf{v}, \mathbf{Y} - \mathbf{X} \rangle)]} \quad [\text{Cauchy-Schwarz}] \\ &\leq e^{2t^2 \sigma^2}. \end{aligned}$$

Therefore, $\langle \mathbf{v}, (\mathbf{Z}(\mathbf{Y}) - \mathbf{X}) \rangle$ is sub-gaussian and satisfies the tail bound

$$\Pr(|\langle \mathbf{v}, (\mathbf{Z}(\mathbf{Y}) - \mathbf{X}) \rangle| > t) \leq 2e^{-t^2/(8\sigma^2)}.$$

Taking $t = c_d/2$, because $\sigma^2 = \sum_{i=0}^d \mathbf{v}_i^2 s_i \leq \max_i s_i \leq \frac{c_d^2/32}{\log(4/\epsilon_d)}$, we have that

$$\Pr(|\langle \mathbf{v}, (\mathbf{Z}(\mathbf{Y}) - \mathbf{X}) \rangle| > \frac{c_d}{2}) \leq 2e^{-c_d^2/(32\sigma^2)} = \epsilon_d/2. \quad (14)$$

Now, we can prove the desired result that

$$\begin{aligned} & \mathbb{E}[(\langle \mathbf{Z}(\mathbf{Y}), \mathbf{v} \rangle)^+] \\ &\geq \mathbb{E} \left[(\langle \mathbf{Z}(\mathbf{Y}), \mathbf{v} \rangle)^+ \mid |\langle \mathbf{v}, (\mathbf{Z}(\mathbf{Y}) - \mathbf{X}) \rangle| \leq \frac{c_d}{2}, \langle \mathbf{X}, \mathbf{v} \rangle \geq c_d \right] \Pr \left(|\langle \mathbf{v}, (\mathbf{Z}(\mathbf{Y}) - \mathbf{X}) \rangle| \leq \frac{c_d}{2}, \langle \mathbf{X}, \mathbf{v} \rangle \geq c_d \right) \\ &\geq \mathbb{E} \left[\frac{c_d}{2} \mid |\langle \mathbf{v}, (\mathbf{Z}(\mathbf{Y}) - \mathbf{X}) \rangle| \leq \frac{c_d}{2}, \langle \mathbf{X}, \mathbf{v} \rangle \geq c_d \right] \Pr \left(|\langle \mathbf{v}, (\mathbf{Z}(\mathbf{Y}) - \mathbf{X}) \rangle| \leq \frac{c_d}{2}, \langle \mathbf{X}, \mathbf{v} \rangle \geq c_d \right) \\ &= \frac{c_d}{2} \Pr \left(|\langle \mathbf{v}, (\mathbf{Z}(\mathbf{Y}) - \mathbf{X}) \rangle| \leq \frac{c_d}{2}, \langle \mathbf{X}, \mathbf{v} \rangle \geq c_d \right) \\ &\geq \frac{c_d}{2} \left(\Pr(\langle \mathbf{X}, \mathbf{v} \rangle \geq c_d) - \Pr \left(|\langle \mathbf{v}, (\mathbf{Z}(\mathbf{Y}) - \mathbf{X}) \rangle| > \frac{c_d}{2} \right) \right) \\ &\geq \frac{c_d}{2} \left(\epsilon_d - \frac{\epsilon_d}{2} \right) \quad [\text{Eq (14) and lemma assum}] \\ &= \frac{c_d \epsilon_d}{4}. \end{aligned} \quad \square$$

H.1 Proof of Lemma A.1

Proof of Lemma A.1. Let B be a Bernoulli random variable such that $\Pr(B = 1) = \epsilon$ and let $Z \sim N(0, 1)$ be independent of B and X . Then we can write $R \sim X \cdot B + Z$.

Then we have that

$$\begin{aligned}
& \mathbb{E}[X \mid R > 0] \\
&= \mathbb{E}[X \mid R > 0, B = 1] \Pr(B = 1 \mid R > 0) + \mathbb{E}[X \mid R > 0, B = 0] \Pr(B = 0 \mid R > 0) \\
&= \mathbb{E}[X \mid X + Z > 0] \Pr(B = 1 \mid R > 0) \\
&= \mathbb{E}[X \mid X + Z > 0] \frac{\Pr(R > 0 \mid B = 1) \Pr(B = 1)}{\Pr(R > 0)} \\
&\geq \mathbb{E}[X \mid X + Z > 0] \Pr(R > 0 \mid B = 1) \Pr(B = 1) \\
&= \mathbb{E}[X \mid X + Z > 0] \Pr(X + Z > 0) \epsilon.
\end{aligned} \tag{15}$$

Next, we need to lower bound $\mathbb{E}[X \mid X + Z > 0] \Pr(X + Z > 0)$. Applying Baye's rule gives

$$\begin{aligned}
& \mathbb{E}[X \mid X + Z > 0] \Pr(X + Z > 0) \\
&= \Pr(X + Z > 0) \int_{-\infty}^{\infty} x d\mu_{X \mid X+Z>0}(x) \\
&= \int_{-\infty}^{\infty} x (\Phi^C(-x)) d\mu_X(x) \\
&= \int_{-\infty}^{\infty} x (\Phi^C(-x) - 1/2) d\mu_X(x) \quad [\mathbb{E}[X] = 0]
\end{aligned}$$

Note that $(\Phi^C(-x) - 1/2)$ has the same sign as x and has magnitude increasing in $|x|$. Therefore,

$$\begin{aligned}
& \geq \int_{x \geq \frac{\sigma_X}{\sqrt{10}}} \frac{\sigma_X}{\sqrt{10}} \mathbb{P}\left(0 \leq Z \leq \frac{\sigma_X}{\sqrt{10}}\right) d\mu_X(x) + \int_{x \leq -\frac{\sigma_X}{\sqrt{10}}} \frac{\sigma_X}{\sqrt{10}} \mathbb{P}\left(0 \leq Z \leq \frac{\sigma_X}{\sqrt{10}}\right) d\mu_X(x) \\
&= \frac{\sigma_X \mathbb{P}(0 \leq Z \leq \frac{\sigma_X}{\sqrt{10}})}{\sqrt{10}} \mathbb{P}(|X| > \frac{\sigma_X}{\sqrt{10}}) \\
&\geq \frac{\sigma_X \mathbb{P}(0 \leq Z \leq \frac{\sigma_X}{\sqrt{10}})}{\sqrt{10}} \left(\frac{4\sigma_X^2}{5K^2 \log\left(\frac{20K^2}{\sigma_X^2}\right)} \right). \quad [\text{Equation (17) below}] \\
&\geq \Omega(\sigma_X^5).
\end{aligned} \tag{16}$$

Combining Equations (15) and (16) gives the desired result of the lemma.

It remains to show the lower bound on $\mathbb{P}(|X| > \frac{\sigma_X}{\sqrt{10}})$ used in the penultimate line above.

Define $a = K^2 \log\left(\frac{20K^2}{\sigma_X^2}\right) \geq \sigma_X^2 \log(10)/2 > \sigma_X^2/10$ (using Equation (4)). Next, we observe

that

$$\begin{aligned}
& \mathbb{E}[X^2] \\
&= \int_0^\infty \mathbb{P}(X^2 > t) dt \\
&= \int_0^\infty \mathbb{P}(X > \sqrt{t}) dt \\
&= \int_0^{\sigma_X^2/10} \mathbb{P}(X > \sqrt{t}) dt + \int_{\sigma_X^2/10}^a \mathbb{P}(X > \sqrt{t}) dt + \int_a^\infty \mathbb{P}(X > \sqrt{t}) dt \\
&\leq \sigma_X^2/10 + \int_{\sigma_X^2/10}^a \mathbb{P}(X > \sqrt{t}) dt + \int_a^\infty 2e^{-t/K^2} dt \\
&= \sigma_X^2/10 + \int_{\sigma_X^2/10}^a \mathbb{P}(X > \sqrt{t}) dt + 2K^2 e^{-a/K^2} \\
&= \sigma_X^2/5 + \int_{\sigma_X^2/10}^a \mathbb{P}(X > \sqrt{t}) dt \quad [\text{Def of } a] \\
&\leq \sigma_X^2/5 + \left(a - \frac{\sigma_X^2}{10}\right) \mathbb{P}\left(X > \frac{\sigma_X}{\sqrt{10}}\right). \quad [\mathbb{P}(X > \sqrt{t}) \text{ monotone decr.}]
\end{aligned}$$

Since $\mathbb{E}[X^2] = \sigma_X^2$, this implies that

$$\left(a - \frac{\sigma_X^2}{10}\right) \mathbb{P}\left(X > \frac{\sigma_X}{\sqrt{10}}\right) \geq \frac{4\sigma_X^2}{5}.$$

Therefore, we can conclude that

$$\mathbb{P}\left(X > \frac{\sigma_X}{\sqrt{10}}\right) \geq \frac{4\sigma_X^2/5}{a - \sigma_X^2/10} \geq \frac{4\sigma_X^2/5}{a} = \frac{4\sigma_X^2}{5K^2 \log\left(\frac{20K^2}{\sigma_X^2}\right)}. \quad (17)$$

By symmetry, identical logic as above gives the desired upper bound on $\mathbb{E}[X \mid R \leq 0]$. \square

I Proof of Proposition 3.2

Proof of Proposition 3.2. We first show that $\mathbf{A}^{(t)}$ on Line 6 satisfies $\mathbf{A}^{(t)} \in S^\perp$ when $\Psi = 1$. Recall \mathbf{x}^* defined as $x_\ell^* = \langle \boldsymbol{\ell}^*, \mathbf{w}_\ell \rangle$ and recall that $\mathbf{z}(\mathbf{y}) = \mathbb{E}[\mathbf{x}^* \mid \hat{\mathbf{y}} = \mathbf{y}]$. By construction,

$$\begin{aligned}
& \mathbb{E}[\mathbf{x}^* \mid \Psi = 1] \\
&= \int \mathbb{E}[\mathbf{x}^* \mid \Psi = 1, \hat{\mathbf{y}} = \mathbf{y}] d\mu_{\hat{\mathbf{y}}|\Psi=1}(\mathbf{y}) \\
&= \int \mathbb{E}[\mathbf{x}^* \mid \hat{\mathbf{y}} = \mathbf{y}] \frac{\Pr(\Psi = 1 \mid \hat{\mathbf{y}} = \mathbf{y})}{\Pr(\Psi = 1)} d\mu_{\hat{\mathbf{y}}}(\mathbf{y}) \quad [\Psi = 1 \text{ is a function of } \hat{\mathbf{y}}] \\
&= \frac{1}{\Pr(\Psi = 1)} \int \mathbb{E}[\mathbf{x}^* \mid \hat{\mathbf{y}} = \mathbf{y}] f(\mathbf{z}(\mathbf{y})) d\mu_{\hat{\mathbf{y}}}(\mathbf{y}) \\
&= \frac{1}{\Pr(\Psi = 1)} \int \mathbf{z}(\mathbf{y}) f(\mathbf{z}(\mathbf{y})) d\mu_{\hat{\mathbf{y}}}(\mathbf{y}) \\
&= \frac{1}{\Pr(\Psi = 1)} \mathbb{E}[\mathbf{z}(\hat{\mathbf{y}}) f(\mathbf{z}(\hat{\mathbf{y}}))] \\
&= 0. \quad [\text{Definition of } f]
\end{aligned}$$

Because $x_\ell^* = \langle \boldsymbol{\ell}^*, \mathbf{w}_\ell \rangle$ for $\ell \leq \ell_\lambda$, this implies that $\mathbb{E}[\langle \boldsymbol{\ell}^*, \mathbf{w}_\ell \rangle \mid \Psi = 1] = 0$ for $\ell \leq \ell_\lambda$. Therefore, we must have that $\mathbb{E}[\boldsymbol{\ell}^* \mid \Psi = 1] \in S^\perp$. By construction of $\mathbf{A}^{(t)}$ in Line 6, this implies that $\mathbf{A}^{(t)} \in S^\perp$ when $\Psi = 1$.

Define

$$\mathbf{A} = \begin{cases} \frac{\mathbb{E}[\boldsymbol{\ell}^* \mid \Psi=1]}{\|\mathbb{E}[\boldsymbol{\ell}^* \mid \Psi=1]\|_2} & \text{if } \mathbb{E}[\boldsymbol{\ell}^* \mid \Psi = 1] \neq 0 \\ \mathbf{w}_{\ell_\lambda+1} & \text{otherwise,} \end{cases}$$

in other words \mathbf{A} is equal to $\mathbf{A}^{(t)}$ when $\Psi = 1$.

By the choice of f , we have that $\mathbb{P}(\Psi = 1) \geq \frac{\epsilon_d c_d}{16 \max(\|\mathbb{E}[\mathbf{z}(\hat{\mathbf{y}})]\|, 1)} \geq \frac{\epsilon_d c_d}{16(K\sqrt{\pi}+1)}$, where in the last line we used that Equation (3) implies

$$\max(\|\mathbb{E}[\mathbf{z}(\hat{\mathbf{y}})]\|, 1) = \max(\|\mathbb{E}[\mathbf{x}^*]\|, 1) \leq \max(\|\mathbb{E}[\boldsymbol{\ell}^*]\|, 1) \leq \max(K\sqrt{\pi}, 1) \leq K\sqrt{\pi} + 1.$$

By construction, we therefore have that for any realization of $\mathbf{z}(\hat{\mathbf{y}})$, the probability that $R = r^{(t)} = \langle \boldsymbol{\ell}^*, \mathbf{A}^{(t)} \rangle + w_t = \langle \boldsymbol{\ell}^*, \mathbf{A} \rangle + w_t$ is exactly $\frac{\epsilon_d c_d}{16(K\sqrt{\pi}+1)}$ and otherwise $R \sim N(0, 1)$.

We can now apply Lemma A.1 with $X = \langle \boldsymbol{\ell}^*, \mathbf{A} \rangle$, and $\epsilon = \frac{\epsilon_d c_d}{16(K\sqrt{\pi}+1)}$ to get that either $\mathbf{a} = \mathbf{w}_{\ell_\lambda+1}$ or

$$|\langle \mathbf{A}, \mathbf{a} \rangle| = |\langle \mathbf{A}, \mathbb{E}[\boldsymbol{\ell}^* \mid 1_{R>0}] \rangle| = |\mathbb{E}[\langle \boldsymbol{\ell}^*, \mathbf{A} \rangle \mid 1_{R>0}]| \geq \frac{c_{L\mathbf{A}.1} \epsilon_d c_d \text{Var}(\langle \boldsymbol{\ell}^*, \mathbf{A} \rangle)^{2.5}}{16(K\sqrt{\pi}+1)} \geq \frac{c_{L\mathbf{A}.1} \epsilon_d c_d c_v^{2.5}}{16(K\sqrt{\pi}+1)},$$

where in the last inequality we used Assumption 2.

Because $\mathbf{A} \in S^\perp$ and $\|\mathbf{A}\| = 1$, the previous equation implies the desired result that $\|\mathcal{P}_{S^\perp}(\mathbf{a})\| \geq \frac{c_{L\mathbf{A}.1} \epsilon_d c_d c_v^{2.5}}{16(K\sqrt{\pi}+1)}$. \square

J Proof of Proposition 3.3

Proof of Proposition 3.3. The first step is to rewrite R from Algorithm 6 Line 7 so that we can apply Lemmas A.2 and A.3.

Define

$$W := \sum_{t'=t}^{t+L-1} \left(w_{t'} - \sum_{k=1}^j (c_k q_k^{t'} - c_k \langle \mathbf{v}_k, \boldsymbol{\ell}^* \rangle) \right) = \sum_{t'=t}^{t+L-1} \left(w_{t'} - \sum_{k=1}^j c_k q_k^{t'} \right) + L \sum_{k=1}^j c_k \langle \mathbf{v}_k, \boldsymbol{\ell}^* \rangle.$$

Note that W is normally distributed with mean 0 and variance $\sigma_W^2 := L(1 + \sum_{k=1}^j c_k^2)$. By construction, we can rewrite R as

$$\begin{aligned} R &= \sum_{t'=t}^{t+L-1} \left(r^{(t')} - \sum_{k=1}^j c_k q_k^{t'} \right) \\ &= \sum_{t'=t}^{t+L-1} \left((\langle \mathbf{a}, \boldsymbol{\ell}^* \rangle + w_{t'}) - \sum_{k=1}^j c_k q_k^{t'} \right) \\ &= \sum_{t'=t}^{t+L-1} \langle \mathbf{a}, \boldsymbol{\ell}^* \rangle - L \sum_{k=1}^j c_k \langle \mathbf{v}_k, \boldsymbol{\ell}^* \rangle + W \\ &= L \langle \mathbf{a}, \boldsymbol{\ell}^* \rangle - L \langle \mathcal{P}_S(\mathbf{a}), \boldsymbol{\ell}^* \rangle + W \quad [\text{Lemma A.5 implies } \mathcal{P}_S(\mathbf{a}) = \sum_{k=1}^j c_k \mathbf{v}_k] \\ &= L \langle (\mathbf{a} - \mathcal{P}_S(\mathbf{a})), \boldsymbol{\ell}^* \rangle + W \\ &= L \langle \mathcal{P}_{S^\perp}(\mathbf{a}), \boldsymbol{\ell}^* \rangle + W \\ &= L \|\mathcal{P}_{S^\perp}(\mathbf{a})\| \langle \mathbf{x}, \boldsymbol{\ell}^* \rangle + W. \end{aligned} \tag{18}$$

Therefore, R is exactly in the form necessary to apply Lemmas A.2 and A.3. In order to apply these lemmas, we need that $\frac{L \|\mathcal{P}_{S^\perp}(\mathbf{a})\|}{\sigma_W} \leq \delta_{L\text{A.2}}$ and $\frac{L \|\mathcal{P}_{S^\perp}(\mathbf{a})\|}{\sigma_W} \leq \delta_{L\text{A.3}}$ respectively.

To see this, note that $\mathcal{P}_{S^\perp}(\mathbf{a}) \leq \sqrt{\lambda}$ (as otherwise ExponentialGrowth would not have been called), and therefore

$$\begin{aligned} \frac{L \|\mathcal{P}_{S^\perp}(\mathbf{a})\|}{\sigma_W} &\leq \frac{L\sqrt{\lambda}}{\sqrt{L(1 + \sum_{k=1}^j c_k^2)}} \\ &= \sqrt{\frac{4\lambda d(\mathbb{E}[\ell^*_{\cdot 1}] + 1)^2}{c_{L\text{A.2}}^2}} \\ &\leq \sqrt{\frac{4\lambda d(K\sqrt{\pi} + 1)^2}{(c_v/\sqrt{8\pi})^2}} \quad [\text{Equation (3), Assum 2}] \\ &\leq \min(\delta_{L\text{A.2}}, \delta_{L\text{A.3}}, 1/c_{L\text{A.3}}), \end{aligned} \tag{19}$$

where in the last line we used $\lambda \leq \min(\delta_{L\text{A.2}}, \delta_{L\text{A.3}}, 1/c_{L\text{A.3}})^2 \frac{(c_v/\sqrt{8\pi})^2}{4d(K\sqrt{\pi}+1)^2}$ by our assumption on λ .

Applying Lemmas A.2 and A.3 gives the following two bounds. Define $\mathbf{y} = \mathbb{E}[\boldsymbol{\ell}^* \mid 1_{R>0}]$. The first is a lower bound on $|\langle \mathbf{x}, \mathbf{y} \rangle|$. Importantly, we can apply Lemma A.2 for $X = \langle \mathbf{x}, \boldsymbol{\ell}^* \rangle$ because of Equations (18) and (19).

$$\begin{aligned}
|\langle \mathbf{x}, \mathbf{y} \rangle| &= |\langle \mathbf{x}, \mathbb{E}[\boldsymbol{\ell}^* \mid 1_{R>0}] \rangle| \\
&= |\mathbb{E}[\langle \mathbf{x}, \boldsymbol{\ell}^* \rangle \mid 1_{R>0}]| \\
&\geq \frac{c_{L\text{A.2}} L \|\mathcal{P}_{S^\perp}(\mathbf{a})\|}{\sigma_W} && \text{[Lemma A.2]} \\
&= \frac{c_{L\text{A.2}} \sqrt{L} \|\mathcal{P}_{S^\perp}(\mathbf{a})\|}{\sqrt{1 + \sum_{k=1}^j c_k^2}} \\
&= c_{L\text{A.2}} \sqrt{\frac{4d(\mathbb{E}[\boldsymbol{\ell}_1^*] + 1)^2}{c_{L\text{A.2}}^2}} \|\mathcal{P}_{S^\perp}(\mathbf{a})\| \\
&= \sqrt{4d(\mathbb{E}[\boldsymbol{\ell}_1^*] + 1)^2} \|\mathcal{P}_{S^\perp}(\mathbf{a})\| \\
&= 2\sqrt{d}(\mathbb{E}[\boldsymbol{\ell}_1^*] + 1) \|\mathcal{P}_{S^\perp}(\mathbf{a})\|. \tag{20}
\end{aligned}$$

The next equation is an upper bound on $\|\mathbf{y}_i\|$ for all $i \in [d]$:

$$\begin{aligned}
|\mathbf{y}_i| &= \mathbb{E}[\boldsymbol{\ell}_i^* \mid 1_{R>0}] \\
&\leq \mathbb{E}[\boldsymbol{\ell}_i^*] + c_{L\text{A.3}} L \|\mathcal{P}_{S^\perp}(\mathbf{a})\| / \sigma_W && \text{[Lemma A.3]} \\
&\leq \mathbb{E}[\boldsymbol{\ell}_i^*] + 1. && \text{[Equation (19)]}
\end{aligned}$$

Using the above equation, we can bound $\|\mathbf{y}\|_2$ as follows. Because $\mathbb{E}[\boldsymbol{\ell}_1^*] \geq \mathbb{E}[\boldsymbol{\ell}_i^*]$ for all i ,

$$\|\mathbf{y}\|_2 \leq \sqrt{\sum_{i=1}^d (\mathbb{E}[\boldsymbol{\ell}_i^*] + 1)^2} \leq \sqrt{d}(\mathbb{E}[\boldsymbol{\ell}_1^*] + 1).$$

Equation (20) implies that $\mathbf{y} \neq \mathbf{0}$. This implies by construction that $\mathbf{b} = \text{Exploit}(1_{R>0}, \mathbf{w}_{\ell_\lambda+1}) = \frac{\mathbf{y}}{\|\mathbf{y}\|}$. Putting everything together, we have that

$$|\langle \mathbf{x}, \mathbf{b} \rangle| = \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|}{\|\mathbf{y}\|_2} \geq \frac{2\sqrt{d}(\mathbb{E}[\boldsymbol{\ell}_1^*] + 1) \|\mathcal{P}_{S^\perp}(\mathbf{a})\|}{\sqrt{d}(\mathbb{E}[\boldsymbol{\ell}_1^*] + 1)} \geq 2 \|\mathcal{P}_{S^\perp}(\mathbf{a})\|.$$

Finally, because $\mathbf{x} \in S^\perp$, this implies the desired result that

$$\|\mathcal{P}_{S^\perp}(\mathbf{b})\| \geq |\langle \mathbf{x}, \mathbf{b} \rangle| \geq 2 \|\mathcal{P}_{S^\perp}(\mathbf{a})\|. \quad \square$$

K Proof of Theorem 3.4

Proof of Theorem 3.4. We begin by proving that Algorithm 4 is BIC. There are four places where we set $\mathbf{A}^{(t)}$. The first is in the Line 4 of Algorithm 4, where we set $\mathbf{A}^{(t)} = \mathbf{e}_1$. This

is BIC because we assumed (without loss of generality) that $\mathbb{E}[\ell^*_i] = 0$ for all $i > 1$ and $\mathbb{E}[\ell^*_1] \geq 0$.

The second place we set $\mathbf{A}^{(t)}$ is in Line 6 of Algorithm 5. This choice of $\mathbf{A}^{(t)}$ is BIC with the signal Ψ by construction and Lemma 2.1.

The third place we set $\mathbf{A}^{(t)}$ is in Line 5 of Algorithm 6. In order for this to be BIC, we must show that every input \mathbf{a} to Algorithm 3 is BIC. The first time Algorithm 6 is used for any fixed value of j , the input action \mathbf{a} is the action returned by Algorithm 5. This is BIC for signal R defined on Line 7 of Algorithm 5 by construction. Each subsequent call to Algorithm 6 for a fixed value of j uses an action \mathbf{a} that is returned by the previous call to Algorithm 6. This is BIC for signal R defined on Line 7 of Algorithm 6.

The final time we set an action is on Line 19 of Algorithm 4. This action is again an action returned by the last call to Algorithm 6, which as argued above is BIC for signal R .

The rest of the proof will focus on bounding the sample complexity of Algorithm 4.

First, we will bound the number of times the inner while loop (Line 14) calls Algorithm 6 for each value of j . By Proposition 3.2, the action returned by Algorithm 5 satisfies $\|\mathcal{P}_{S^\perp}(\mathbf{a})\| \geq c_{P3.2} c_v^{2.5} \epsilon_d c_d$. Furthermore, by Proposition 3.3, $\|\mathcal{P}_{S^\perp}(\mathbf{a})\|$ doubles with each call to Algorithm 6. Therefore, $\|\mathcal{P}_{S^\perp}(\mathbf{a})\| \geq \sqrt{\lambda}$ will be satisfied after at most $\log_2 \left(\frac{\sqrt{\lambda}}{c_{P3.2} c_v^{2.5} \epsilon_d c_d} \right) = O \left(\log \left(\frac{1}{c_v \epsilon_d c_d} \right) \right)$ calls to Algorithm 6.

Next we will bound the number of steps in each call to Algorithm 6, which is equivalent to bounding the L defined on Line 4 of Algorithm 6. To do this, we note that the c_i in Algorithm 6 are the same as the c_i in Lemma A.5 with $\epsilon = \lambda$, $\ell = \ell_\lambda$, $\mathbf{u} = \mathcal{P}_{S^\perp}(\mathbf{a})$, and $\mathbf{v}_1, \dots, \mathbf{v}_j$. This implies that

$$\sum_{k=1}^j c_k^2 \leq \frac{1}{\lambda}. \quad [\text{Lemma A.5}]$$

Therefore, we can bound L as follows:

$$\begin{aligned} L &= \frac{4d(\mathbb{E}[\ell^*_1] + 1)(1 + \sum_{k=1}^j c_k^2)}{c_{LA.2}^2} \leq \frac{4d(\mathbb{E}[\ell^*_1] + 1)(1 + \frac{1}{\lambda})}{c_{LA.2}^2} \\ &\leq \frac{4d(K\sqrt{\pi} + 1)(1 + \frac{1}{\lambda})}{c_v^2 / (8\pi)} \quad [\text{Assum 2, Eq (3)}] \quad (21) \\ &= O \left(\frac{d}{\lambda c_v^2} \right). \end{aligned}$$

For each loop of the while loop on Line 9, we also have $\kappa = O(\frac{\log(1/\epsilon_d)}{\lambda c_d^2} + \frac{d}{\lambda c_v^2})$ steps in the loop on Line 18. All together, this gives that each iteration of the loop on Line 9 takes at most

$$O \left(\frac{d \log(\frac{1}{c_v \epsilon_d c_d})}{\lambda c_v^2} + \frac{\log(1/\epsilon_d)}{\lambda c_d^2} \right) = O \left(\log \left(\frac{1}{c_v \epsilon_d c_d} \right) \left(\frac{d}{\lambda c_v^2} + \frac{1}{\lambda c_d^2} \right) \right) \quad (22)$$

steps. Next, we will bound the number of iterations of the while loop on Line 9.

For each j , we will apply Lemma 2.3 with $\epsilon = \lambda$, $\mathbf{u} = \mathbf{v}_{j+1}$, and the vectors $\mathbf{v}_1, \dots, \mathbf{v}_j$. By construction of the algorithm, S^\perp is non-empty because the algorithm has not yet terminated, and $\|\mathcal{P}_{S^\perp}(\mathbf{v}_{j+1})\|^2 \geq \lambda$ by the termination condition of the while loop on Line 14 of Algorithm 4. Therefore, this satisfies the assumption of Lemma 2.3. Define $\lambda_1^j, \dots, \lambda_d^j$ as the eigenvalues of $\mathbf{M}^j := \sum_{i=1}^j \mathbf{v}_i^{\otimes 2}$ and define ℓ^j as the largest index such that $\lambda_{\ell^j}^j \geq 200d^3/\lambda^2$ (and $\ell^j = 0$ if all eigenvalues of \mathbf{M}^j are less than $200d^3/\lambda^2$). Now define

$$\Delta^j = \sum_{i=\ell^j+1}^d \left(\frac{200d^3}{\lambda^2} - \lambda_i \right).$$

Note that for any fixed i , the i th eigenvalue does not decrease between \mathbf{M}^j and \mathbf{M}^{j+1} . Because of this monotonicity, Lemma 2.3 implies that for every round j , either

$$\ell^{j+1} \geq \ell^j + 1 \quad \text{or} \quad \Delta^{j+1} \leq \Delta^j - \frac{\lambda}{2}.$$

Because $\ell^1 \geq 0$ and $\Delta^1 \leq \frac{200d^3}{\lambda^2} \cdot d = \frac{200d^4}{\lambda^2}$, this implies that after $\frac{200d^4}{\lambda/2} + d$ applications of Lemma 2.3, either $\ell^j = d$ or $\Delta^j = 0$. This means that after $\frac{400d^4}{\lambda^2} + d$ applications of Lemma 2.3, the smallest eigenvalue of \mathbf{M}^j must be at least $200d^3/\lambda^2 \geq \lambda$. However, this means that the algorithm must terminate before round $400d^4/\lambda^3 + d$. Therefore, the number of iterations of the while loop on Line 9 is less than $O(d^4/\lambda^3)$. Putting everything together, the total number of steps needed for λ -exploration is upper bounded by

$$O\left(\log\left(\frac{1}{c_v \epsilon_d c_d}\right) \left(\frac{d}{\lambda c_v^2} + \frac{1}{\lambda c_d^2}\right)\right) \cdot O\left(\frac{d^4}{\lambda^3}\right) = O\left(\log\left(\frac{1}{c_v \epsilon_d c_d}\right) \left(\frac{d^5}{\lambda^4 c_v^2} + \frac{d^4}{c_d^2 \lambda^4}\right)\right). \quad \square$$

L Proof of Proposition 1.4

Proof of Proposition 1.4. First, for any unit vector \mathbf{v} , we have $B_{r/3}(2r\mathbf{v}/3) \subseteq \mathcal{K} \subseteq B_1(0)$. Therefore

$$\mu(B_{r/3}(2r\mathbf{v}/3)) = \text{Vol}(B_{r/3}(2r\mathbf{v}/3))/\text{Vol}(\mathcal{K}) \geq \text{Vol}(B_{r/3}(2r\mathbf{v}/3))/\text{Vol}(B_1(0)) = (r/3)^d.$$

Since $\langle \mathbf{x}, \mathbf{v} \rangle \geq r/3$ for all $\mathbf{x} \in B_{r/3}(2r\mathbf{v}/3)$, this confirms the values $(c_d, \epsilon_d) = (r/3, (r/3)^d)$.

The bound on c_v follows by [Sel23, Lemma 3.2] and Jensen's inequality since \mathcal{K} has width at least $2r$ in any direction. The bound on K is trivial since $2e^{-(t/1.25)^2} \geq 1$ for $|t| \leq 1$. \square

M Proof of Proposition 1.5

We first recall several useful facts on log-concave distributions. Throughout we take μ to be α -log-concave and β -log-smooth with mode \mathbf{x}^* and mean $\bar{\mathbf{x}}$, possibly in dimension 1. (The proof will use 1-dimensional projections of the original measure μ .) We will write $\mathbf{x} \sim \mu$ instead of $\ell^* \sim \mu$.

Fact M.1 ([DCWY19, Lemma 5], [DM19, Theorem 1]). For $x \sim \mu$, we have $\mathbb{E}[\|x - x^*\|^2] \leq 1/\alpha$ and with probability $1 - \delta$:

$$\|x - x^*\|_2 \leq 2\alpha^{-1/2} \left(1 + \sqrt{\frac{\log(1/\delta)}{d}} + \sqrt[4]{\frac{\log(1/\delta)}{d}} \right). \quad (23)$$

Fact M.2 ([CP23, Lemma 2]). We have the covariance bounds

$$\frac{I_d}{\alpha d} \succeq \text{Cov}(\mu) \succeq \frac{I_d}{\beta d}. \quad (24)$$

Fact M.3. Any 1-dimensional projection of μ is also αd -log-concave and βd -log-smooth.

Proof. Preservation of strong log-concavity under projection is well known, see e.g. [SW14, Theorem 3.8]. For log-smoothness, supposing for convenience that the projection is onto the first coordinate axis, the claim is proved by the following standard computation. With $e^{-f(\mathbf{x})}$ the density of μ and $e^{-g(x)}$ the density of the projection of μ to the first coordinate axis, one may compute as in [SW14, Proof of Proposition 7.1] that

$$g''(x) = \mathbb{E}^\mu[\partial_{1,1}f(\mathbf{x})|x_1 = x] - \text{Var}^\mu[\partial_{1,1}f(\mathbf{x})|x_1 = x] \leq \mathbb{E}^\mu[\partial_{1,1}f(\mathbf{x})|x_1 = x] \leq \beta d.$$

This completes the proof. \square

Proof of Proposition 1.5. We have $c_v \geq \frac{1}{\beta d}$ directly from (24).

For ϵ_d , let $\bar{\mathbf{x}}$ be the mean under μ and note that from (24), we find

$$\begin{aligned} \|\bar{\mathbf{x}} - \mathbf{x}^*\| &= \sup_{\|\mathbf{w}\|=1} \langle \bar{\mathbf{x}} - \mathbf{x}^*, \mathbf{w} \rangle \\ &= \sup_{\|\mathbf{w}\|=1} \mathbb{E}^{\mathbf{x} \sim \mu}[\langle \mathbf{x} - \mathbf{x}^*, \mathbf{w} \rangle] \\ &\leq \sqrt{\sup_{\|\mathbf{w}\|=1} \mathbb{E}^{\mathbf{x} \sim \mu}[\langle \mathbf{x} - \mathbf{x}^*, \mathbf{w} \rangle^2]} \\ &\leq \sqrt{\langle \text{Cov}(\mu), \mathbf{w}^{\otimes 2} \rangle} \\ &\leq 1/\sqrt{\alpha d}. \end{aligned}$$

Fixing a unit vector \mathbf{v} as in Assumption 1.2, we consider the projection P onto the 1-dimensional subspace spanned by \mathbf{v} , and let $P(\mu)$ be the pushforward of μ under the projection (to which Fact M.3 applies). Identifying $P(\mathbb{R}^d)$ isometrically with \mathbb{R} , let \hat{x} be the mode of $P(\mu)$. Then the same argument as above applies to $P(\mu)$ shows $\|P(\bar{\mathbf{x}}) - \hat{x}\| \leq 1/\sqrt{\alpha d}$, and so

$$\|\hat{x}\| \leq \|\mathbf{x}^*\| + \frac{2}{\sqrt{\alpha d}} \leq \gamma + \frac{2}{\sqrt{\alpha d}}.$$

(Note that if $\bar{\mathbf{x}} = 0$ then this shows $\|\hat{x}\| \leq 1/\sqrt{\alpha d}$, which following the arguments below leads to $\epsilon_d \geq \Omega(1)$ as mentioned below Proposition 1.5.)

Write $f : \mathbb{R} \rightarrow \mathbb{R}_+$ for the density of $P(\mu)$, and $g(x) = \log f(x)$. We have $f'(\hat{x}) = 0$ and so $g'(\hat{x}) = 0$ also. By Fact M.3, we have $g''(x) \in [-\beta d, -\alpha d]$ for all x , so for $x \geq \hat{x}$ we have:

$$g'(x) = g'(x) - g'(\hat{x}) = \int_{\hat{x}}^x g''(y) dy \in [-\beta d(x - \hat{x}), -\alpha d(x - \hat{x})].$$

Integrating again, we find

$$g(x) - g(\hat{x}) = \int_{\hat{x}}^x g'(y) dy \in [-\beta d(x - \hat{x})^2/2, -\alpha d(x - \hat{x})^2/2].$$

Identical reasoning gives the same conclusion for $x \leq \hat{x}$. Translating back to $f = e^g$, we conclude that for each $x \in \mathbb{R}$:

$$e^{-\beta d|x-\hat{x}|^2/2} \leq \frac{f(x)}{f(\hat{x})} \leq e^{-\alpha d|x-\hat{x}|^2/2}.$$

It follows that for $\mathbf{x} = \ell^* \sim \mu$ and $x \sim P(\mu)$:

$$\Pr[\langle \mathbf{v}, \mathbf{x} \rangle \geq c_d] \geq \Pr[x \geq \gamma + \frac{2}{\sqrt{\alpha d}} + c_d].$$

Letting $J = \gamma + \frac{2}{\sqrt{\alpha d}} + c_d$, the latter probability is at least

$$\frac{\int_J^\infty e^{-\beta dz^2/2} dz}{\int_{\mathbb{R}} e^{-\alpha dz^2/2} dz} = \sqrt{\alpha/\beta} \cdot \Phi^C(J\sqrt{\beta d}) \geq \frac{J e^{-J^2 \beta d/2} \sqrt{\alpha d}}{(1 + J^2 \beta d) \sqrt{2\pi}}.$$

The last inequality follows from the classical bound $\Phi^C(\kappa) \geq \frac{\varphi(\kappa)\kappa}{1+\kappa^2}$ where φ is the standard Gaussian density [Gor41]. This confirms the value of ϵ_d .

For K we consider a similar projection, and note that by Fact M.3 and the Bakry-Emery theory for strongly log-concave measures (see e.g. [AGZ10, Lemma 2.3.3]), we have

$$\mathbb{E}[e^{\lambda \langle \mathbf{v}, \mathbf{x} - \bar{\mathbf{x}} \rangle}] \leq e^{\lambda^2 \alpha d/2}, \quad \forall \lambda \in \mathbb{R}.$$

Thus with $J_0 = \gamma + \frac{1}{\sqrt{\alpha d}} \geq \|\langle \bar{\mathbf{x}}, \mathbf{v} \rangle\|$, we have (using $\lambda J_0 \leq \frac{1+\lambda^2 J_0^2}{2}$):

$$\mathbb{E}[e^{\lambda \langle \mathbf{v}, \mathbf{x} \rangle}] \leq e^{(\lambda^2 \alpha d/2) + \lambda J_0} \leq e^{0.5} \cdot e^{\lambda^2 (J_0^2 + \alpha d)/2} \leq 2e^{\lambda^2 (J_0^2 + \alpha d)/2}.$$

It follows by the usual Markov inequality arguments that

$$\Pr[|\langle \mathbf{v}, \mathbf{x} \rangle| \geq t] \leq 4e^{-\frac{t^2}{2(J_0^2 + \alpha d)}}.$$

Since probabilities are at most 1 and $a \leq \sqrt{a}$ for $a \leq 1$ we find

$$\Pr[|\langle \mathbf{v}, \mathbf{x} \rangle| \geq t] \leq 2e^{-\frac{t^2}{4(J_0^2 + \alpha d)}}$$

which completes the verification of K since $(J_0^2 + \alpha d)^{1/2} \leq J_0 + \sqrt{\alpha d}$.

For the counterexample, we may take $(\alpha, \beta) = (1, 2)$ and let ν be the distribution on \mathbb{R} with density proportional to $e^{-dx^2 \cdot (1+1_{x \geq 0})/2}$. Then let $\mu = \nu^{\otimes d}$, so that $x \sim \mu$ has IID coordinates with law ν . Then the mode \mathbf{x}^* is indeed zero but the mean of ν is non-zero, so taking $\mathbf{v} = (1, 1, \dots, 1)/\sqrt{d}$, a Chernoff estimate shows $\Pr^{\mathbf{x} \sim \mu}[\langle \mathbf{x}, \mathbf{v} \rangle \geq 0] \leq e^{-\Omega(d)}$. \square