A Reduction from Linear Contextual Bandits Lower Bounds to Estimations Lower Bounds

Jiahao He [1] **Jiheng Zhang** [1] **Rachel Q. Zhang** [1]

Abstract

Linear contextual bandits and their variants are usually solved using algorithms guided by parameter estimation. Cauchy-Schwartz inequality established that estimation errors dominate algorithm regrets, and thus, accurate estimators suffice to guarantee algorithms with low regrets. In this paper, we complete the reverse direction by establishing the necessity. In particular, we provide a generic transformation from algorithms for linear contextual bandits to estimators for linear models, and show that algorithm regrets dominate estimation errors of their induced estimators, i.e., low-regret algorithms must imply accurate estimators. Moreover, our analysis reduces the regret lower bound to an estimation error, bridging the lower bound analysis in linear contextual bandit problems and linear regression.

1. Introduction

Contextual bandit is an extension of the multi-armed bandit problems that incorporates individual information, i.e., *context*. In most effective algorithms for linear contextual bandits and their variants developed in the literature, e.g., (Auer, 2002; Li et al., 2010; Goldenshluger & Zeevi, 2013; Kim & Paik, 2019; Han et al., 2021), actions are guided by one or a set of estimators (e.g., OLS or LASSO) for the parameters in the linear reward function. Cauchy-Schwartz inequality establishes that estimation errors dominate regrets. Thus, estimators with small errors suffice to guarantee algorithms with low regrets. However, one naturally suspects their necessity: as the action space is much smaller than the parameter space, is learning the entire reward function indeed necessary for learning the optimal action?

[*]Equal contribution [1]Department of Industrial Engineering and Decision Analytics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China. Correspondence to: Jiahao He <jhebe@connect.ust.hk>.

Proceedings of the 39^{th} International Conference on Machine Learning, Baltimore, Maryland, USA, PMLR 162, 2022.

In this paper, we provide an affirmative answer to this question. That is, developing low-regret algorithms is essentially a procedure of finding accurate estimations. Along with Cauchy-Schwartz inequality, we complete the equivalence between stochastic contextual bandit algorithms and the estimation of the reward function. We construct an *algorithm-based estimator* for any given algorithm of a stochastic linear contextual bandit problem (or its variants), opposite to developing *estimator-based algorithms* in the literature. We show that the regret of the given algorithm dominates the estimation error of the constructed estimator. Our construction and analysis remain valid under additional constraints (e.g., privacy, batch) or structures (e.g., sparsity).

As a byproduct, our work provides a principled approach to establishing a regret lower bound (characterizing the difficulty) of linear contextual bandit problems by reducing the regret to an estimation error. In the literature, lower bounds for linear contextual bandit problems are usually established via specific constructions of hard problem instances. With our reduction, we can effortlessly obtain a regret lower bound by applying existing minimax bounds and constructions in estimation theory, which has been well studied in the literature, e.g., (Duchi & Wainwright, 2013; Duchi, 2016; Duchi et al., 2018; Wang & Xu, 2019; Acharya et al., 2021). Under this principle, we revisit some established lower bounds in the literature and derive some new lower bounds for various stochastic linear contextual bandit problems.

1.1. Related Work

In the literature, contexts can be generated either from an i.i.d. distribution (*stochastic* contexts) or any arbitrary procedure (*adversarial* contexts). Since a lower (upper) bound under stochastic (adversarial) contexts is also a lower (upper) bound under adversarial (stochastic) contexts, here we review literature with both types of contexts. Let T be the total number of periods, K be the total number of arms, and d be dimension of context vector.

The reward functions can be identical for all arms (*single parameter setting*), or arm-dependent (*multiple parameter setting*). Kannan et al. (2018) emphasized that a problem under the multiple parameter setting with dimension d is equivalent to that under the single parameter setting with dimension Kd.

The Basic Model. Under adversarial contexts and single parameter setting, Auer (2002), Chu et al. (2011), and Li et al. (2010) provide upper bounds of order $O\left(\sqrt{dT}\log^{3/2}(KT/\delta)\right)$ [1], $O\left(\sqrt{Td}\log^{3/2}(KT\ln(T)/\delta)\right)$, $O\left(\sqrt{dT\log(T)\log(K)}\left(\log(\log(dT))\right)^{\gamma}\right)$, respectively, for various variants of UCB-type algorithms. The latter two also provide lower bounds of order $\Omega\left(\sqrt{dT}\right)$ and $\Omega\left(\sqrt{dT\log(T/d)\log(K)}\right)$, respectively.

For stochastic contexts, with multiple parameters, Goldenshluger & Zeevi (2013) provided an $O\left(d^3\log(T)\right)$ upper bound for an OLS-based algorithm in 2-armed problems. They also provided an $\Omega\left(\log(T)\right)$ lower bound, which can be further improved to $\Omega\left(d\log(T)\right)$ according to (Bastani & Bayati, 2020). With an additional assumption on the diversity of the random context, Bastani et al. (2021) showed that even a greedy algorithm can guarantee an upper bound of the order $O\left(d^3\log^{3/2}(d)\log(T)\right)$. Kannan et al. (2018) demonstrated that all the order optimality of greedy algorithms resulting from the merit of stochastic contexts can be extended to adversarial contexts perturbed by Gaussian noise.

Sparsity. When d is much larger than T, the aforementioned bounds are linear in T and not informative. In the sparse setting, only s_0 number of elements in the parameter vector are non-zero. With multiple parameters, Bastani & Bayati (2020) developed a LASSO-based algorithm and provided an $O\left(s_0^2K\left(\log(T)+\log(d)\right)^2\right)$ regret upper bound under several technical assumptions including the existence of a probabilistic constant gap between the rewards of the optimal and suboptimal arms. They also outlined a proof that implies a regret lower bound of $\Omega\left(s_0\log(T)\right)$ when $K=2$. With a single parameter, Kim & Paik (2019) proposed an algorithm based on doubly-robust technique that guarantees an $O\left(s_0\log(dT)\sqrt{T}\right)$ upper bound without the probabilistic gap assumption when $K=2$. Oh et al. (2021) proposed an algorithm that does not require prior knowledge of s_0 and guarantees an $O\left(s_0\sqrt{T\log(dT)}\right)$ upper bound for $K=2$. Wang & Xu (2019) also studied sub-Gaussian contexts and developed an algorithm with an $\tilde{O}\left(\sqrt{s_0T}\right)$ regret upper bound under the assumption that all coordinates of the context are independent. Li et al. (2021) proposed several general notions of sparsity in high-dimensional stochastic linear contextual bandit problems, including low rank structures and group sparsity when the contexts are treated as matrices. They also provided a unified algorithm and framework to establish regret upper bounds for these problems.

[1]For non-negative functions f,g, $f(x)=O(g(x)) \iff g(x)=\Omega(f(x)) \iff \forall x, f(x)\le cg(x)$ for some constant $c>0$ independent of x. $\tilde{O}(\cdot)$ and $\tilde{\Omega}(\cdot)$ denote the respective meanings within multiplicative logarithmic factors.

With Privacy Constraints. *Differential privacy* is a well-recognized constraint of privacy preservation in data analysis. However, Shariff & Sheffet (2018) showed enforcing the constraint leads to $\Omega(T)$ regret and thus introduced a relaxed notion of (ε,δ)*-joint differential privacy* and developed a privacy-preserving algorithm that achieves a regret upper bound of $\tilde{O}\left(d^{3/4}\sqrt{T/\varepsilon}\right)$, for the single parameter setting with adversarial contexts. Alternatively, Zheng et al. (2020) introduced the notion of *locally differentially privacy* and developed an algorithm that achieves a regret upper bound of $\tilde{O}\left((dT)^{3/4}/\varepsilon\right)$ under adversarial contexts. Han et al. (2021) provided an SGD-based algorithm with a regret upper bound of $\tilde{O}\left(\sqrt{dT}/\varepsilon\right)$ and a lower bound of $\Omega\left(\sqrt{dT}/\varepsilon\right)$ under locally differential privacy and stochastic contexts. Their algorithm can also be extended to problems with generalized linear rewards.

With Batch Constraints. Collecting information about an actual reward or updating an estimator is often costly. As such, some studies have considered the case in which reward information is collected in batches. Given a fixed number of M batches, Han et al. (2020) developed algorithms that achieve a regret upper bound of $\tilde{O}\left(\sqrt{dT}+dT/M\right)$ and a lower bound of $\Omega\left(\sqrt{dT}+\min\left\{\frac{T\sqrt{d}}{M},\frac{T}{\sqrt{M}}\right\}\right)$ under adversarial contexts, and a regret upper bound of $\tilde{O}\left(\sqrt{dT}\left(\frac{T}{d^2}\right)^{\frac{1}{2^{M+1}-2}}\right)$ and a lower bound of $\Omega\left(\sqrt{dT}\left(\frac{T}{d^2}\right)^{\frac{1}{2^{M+1}-2}}\right)$ under stochastic contexts when $K=2$. Ren & Zhou (2020) studied problems with high dimensional stochastic contexts and a sparse parameter vector in the reward function. By allowing the dynamic adjustment of batch sizes, they developed an algorithm with a regret upper bound of $\tilde{O}\left(\sqrt{s_0T}\left(\frac{T}{s_0}\right)^{\frac{1}{2^{M+1}-2}}\right)$ and established a lower bound of $\Omega\left(\max\left\{\frac{\sqrt{s_0T}}{M^2}\left(\frac{T}{s_0}\right)^{\frac{1}{2^{M+1}-2}},\sqrt{s_0T}\right\}\right)$ when $K=2$. Note that these problems reduce to the basic model when $M=T$ and all the bounds apply to the basic model under this case.

1.2. Comparison

Note that our lower bounds are established for problems with stochastic contexts, which are naturally lower bounds for problems with adversarial contexts. In Table 1, we compare lower bounds derived using our proposed technique with existing upper and lower bounds. Empty cells in the table indicate absence in the literature

Model	Upper Bound from Literature	Lower Bound from Literature	Our Lower Bound
Without a margin condition	$\tilde{O}(\sqrt{Td})$	$\Omega(dT)$	$\Omega(\sqrt{dT\log K})$
With a margin condition	$O(d^3\log T)$	$\Omega(d\log T)$	$\Omega(d\log(\frac{T\log K}{d}))$
With a batch constraint (stochastic context)	$O\left(\sqrt{dT}\left(\frac{T}{d^2}\right)^{\frac{1}{2(2^M-1)}}\right)$	$\Omega\left(\sqrt{dT}\left(\frac{T}{d^2}\right)^{\frac{1}{2(2^M-1)}}\right)$	$\Omega\left(\sqrt{dT}\left(\frac{T}{d^2}\right)^{\frac{1}{2(2^M-1)}}\right)$
Sparse	$O(\sqrt{s_0T\log(dT)})$	$\Omega(\sqrt{s_0T})$	$\Omega(\sqrt{s_0T\log K\log(d/s_0)})$
Jointly differentially private	$O(d^{3/4}\sqrt{T}/\sqrt{\varepsilon})$ for adversarial context $\tilde{O}(\sqrt{dT}+\frac{d\log(1/\delta)}{\varepsilon})$ can be achieved (absent from literature)		$\Omega(\sqrt{dT\log K}+\frac{d}{\varepsilon+\delta})$
Locally differentially private with sparsity			$\frac{\sqrt{dT}\log d}{\varepsilon}$

Table 1. Comparison

2. A Generic Formulation

Consider K-armed linear contextual bandit problems represented by $(\boldsymbol{\Theta}^*, \mathcal{F}_K)$ where $\boldsymbol{\Theta}^* \subseteq \mathbb{R}^d$ and $\mathcal{F}_K$ is a set of $d \times K$ dimensional distributions. A problem instance is described by $(\boldsymbol{\theta}^*, F) \in (\boldsymbol{\Theta}^*, \mathcal{F}_K)$ where $\boldsymbol{\theta}^*$ is the unknown parameter predicative of the rewards. At the beginning of period t, the decision maker observes the context $\boldsymbol{X}_t = (\boldsymbol{x}_{t1}, \ldots, \boldsymbol{x}_{tK})$ independently drawn from the distribution $F \in \mathcal{F}_K$ over the time, where $\boldsymbol{x}_{ta} \in \mathbb{R}^d$ is the context related to arm $a \in \{1, \ldots, K\}$. An algorithm $\pi \in \Pi$ selects an arm $a_t^\pi \in \{1, \ldots, K\}$ based on the current context $\boldsymbol{X}_t$ and the historical information up to the end of period $t-1$, denoted as $\mathcal{H}_t$. Once the arm a_t^π is pulled, the decision maker obtains the linear reward $y_{ta_t^\pi} = \boldsymbol{x}_{ta_t^\pi}^\top \boldsymbol{\theta}^* + \xi_{ta_t^\pi}$, where ξ_{ta}, $1 \le a \le K$ and $t = 1, 2, \ldots$, are independent standard normal random noises.

The effectiveness of algorithm π for a single incidence $(\boldsymbol{\theta}^*, F)$ is measured by the accumulative expected regret of π over T periods: $R_{[T]}^\pi(\boldsymbol{\theta}^*, F) = \sum_{t=1}^{T} E\left[R_t^\pi(\boldsymbol{\theta}^*, F)\right]$, where

$$R_t^\pi(\boldsymbol{\theta}^*, F) = \max_{1\le a\le K}\left\{\boldsymbol{x}_{ta}^\top\boldsymbol{\theta}^*\right\} - \boldsymbol{x}_{t,a_t^\pi}^\top\boldsymbol{\theta}^*$$

is the regret of π in period t. The effectiveness of algorithm π for the set of problems $(\boldsymbol{\Theta}^*, \mathcal{F}_K)$ is measured by its expected worst-case regret

$$R_{[T]}^\pi(\boldsymbol{\Theta}^*, \mathcal{F}_K) = \sup_{(\boldsymbol{\theta}^*, F)\in(\boldsymbol{\Theta}^*, \mathcal{F}_K)} R_{[T]}^\pi\left(\boldsymbol{\theta}^*, F\right).$$

The minimum attainable worst-case regret

$$R_{[T]}(\boldsymbol{\Theta}^*, \mathcal{F}_K) = \inf_{\pi\in\Pi} R_{[T]}^\pi(\boldsymbol{\Theta}^*, \mathcal{F}_K)$$

is the regret of the 'best' algorithm and measures the difficulty of the set of the problems $(\boldsymbol{\Theta}^*, \mathcal{F}_K)$.

A deep understanding of $R_{[T]}(\boldsymbol{\Theta}^*, \mathcal{F}_K)$ can help identify the obstacle in developing effective algorithms and provide a benchmark for developed algorithms. There have been abundant studies in the literature to develop various algorithms and provide upper bounds of $R_{[T]}^\pi(\boldsymbol{\Theta}^*, \mathcal{F}_K)$ to establish the effectiveness of an developed algorithm $\pi \in \Pi$. Note that any upper bound for an algorithm $\pi \in \Pi$ naturally serve as an *upper bound* of $R_{[T]}(\boldsymbol{\Theta}^*, \mathcal{F}_K)$. Our focus is to find a tight *lower bound* of $R_{[T]}(\boldsymbol{\Theta}^*, \mathcal{F}_K)$, which shall take the set Π into consideration, rather than considering a single algorithm.

Our methods and result can accommodate several variants of linear contextual bandit problems in the literature by specifying $\boldsymbol{\Theta}^*, \mathcal{F}_K, \mathcal{H}_t$, or Π as follows.

Sparsity. Sparsity can be described by specific choice of $\boldsymbol{\Theta}^*$. For instance, $\boldsymbol{\Theta}^* = \{\boldsymbol{\theta}^* : \|\boldsymbol{\theta}^*\|_0 = s_0, \|\boldsymbol{\theta}^*\|_2 \le 1\}$ for some $s_0 \ll d$ describes sparse bandits, i.e., vectors in $\boldsymbol{\Theta}^*$ have exactly s_0 number of non-zero coordinates. In low-rank bandits where $d = d_1 \times d_2$, as introduced in (Li et al., 2021), $\boldsymbol{\Theta}^*$ consists of all $\boldsymbol{\theta}^* \in \mathbb{R}^d = \mathbb{R}^{d_1\times d_2}$ such that, when treated as a $d_1 \times d_2$ matrix, $\text{rank}(\boldsymbol{\theta}^*) \le d_0$ for some $d_0 \ll \min\{d_1, d_2\}$.

Batch Constraint. Batch constraint can be described by specific $\mathcal{H}_t$. Let $1 \le t_1 < t_2 < \cdots < t_M = T$. If rewards are revealed in batches at the end of period t_j, $j = 1, \ldots, M$, then for $t \in [t_j + 1, t_{j+1}]$ the history $\mathcal{H}_t = \{(\boldsymbol{X}_s, \boldsymbol{x}_{s,a_s^\pi}, y_{s,a_s^\pi}) : 1 \le s \le t_j\}$. If rewards are revealed in each period, i.e., the number of batches is the same as T, then $\mathcal{H}_t = \left\{\left(\boldsymbol{X}_s, \boldsymbol{x}_{s,a_s^\pi}, y_{s,a_s^\pi}\right) : 1 \le s \le t-1\right\}$.

Privacy Constraint. Privacy constraints can be described by specific Π. Here we provide two common notions of privacy.

1. (ε, δ)-jointly differentially private algorithms. Let $\mathcal{H}_t = \{(\boldsymbol{X}_s, \boldsymbol{x}_{s,a_s^\pi}, y_{s,a_s^\pi}) : s = 1, \ldots, t\}$ and $\mathcal{H}_t'$ be its *adjacent* history, i.e., $\mathcal{H}_t$ and $\mathcal{H}_t'$ are same for all $s \in \{1, \ldots, t-1\}$ except one. We can define Π to be the set of all algorithms satisfying that for any sequence of future actions $a_{[t:T]} := (a_{t+1}, \ldots, a_T)$,

$$P(a_{[t:T]}^\pi = a_{[t:T]}|\mathcal{H}_t) \le e^\varepsilon P(a_{[t:T]}^\pi = a_{[t:T]}|\mathcal{H}_t') + \delta.$$

2. ε-locally differentially private algorithms under a_t^π only depends on $\boldsymbol{X}_t$ and $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_t)$, where $\boldsymbol{w}_t$ is a private view of data $\left(\boldsymbol{X}_t, y_{t,a_t^\pi}\right)$ satisfying (a) $\boldsymbol{w}_t$ is independent of $\{(\boldsymbol{X}_s, y_{s,a_s^\pi}, \boldsymbol{w}_s) : 1 \leq s < t\}$, conditioned on the current observations $\{(\boldsymbol{X}_t, y_{t,a_t^\pi})\}$ and past private samples $\{\boldsymbol{w}_s : 1 \leq s < t\}$; (b) for any measurable set W,

$$\sup_{\boldsymbol{u},\boldsymbol{v} \in \mathbb{R}^{dK+1}} \frac{P\left(\boldsymbol{w}_t \in W \middle| (\boldsymbol{X}_t, y_{t,a_t^\pi}) = \boldsymbol{u}\right)}{P\left(\boldsymbol{w}_t \in W \middle| (\boldsymbol{X}_t, y_{t,a_t^\pi}) = \boldsymbol{v}\right)} \leq e^{\varepsilon}.$$

In the literature, $\mathcal{F}_K$ is usually the set of $d \times K$ dimensional sub-Gaussian distributions or distributions with bounded supports. Combinations of the above specifications of $(\boldsymbol{\Theta}^*, \mathcal{H}_t, \Pi)$ describe various problems, e.g., sparse bandits under batch constraints, private bandits under batch constraints, sparse bandits under privacy constraints.

3. A Generic Lower Bound

A lower bound of $R_{[T]}(\boldsymbol{\Theta}^*, \mathcal{F}_K)$ is usually established by analyzing an instance $(\boldsymbol{\theta}^*, F)$ such that (a) $R_{[T]}(\boldsymbol{\theta}^*, F)$ is likely to be large, (b) a tight and tractable lower bound of $R_{[T]}^\pi(\boldsymbol{\theta}^*, F)$ is available for any algorithm $\pi \in \Pi$. We start with the 2-armed problems in Section 3.1 and then extend to the K-armed problems in Section 3.2. In Section 3.3, we demonstrate the tightness of the derived bounds by comparing them with a regret upper bound for a greedy algorithm.

3.1. The 2-Armed Problems

With two arms, an algorithm in each period basically classifies the contexts into two classes, each consisting of contexts under which the algorithm will select the same arm. Note that the optimal classification which leads to 0 regret is a linear one: one should choose arm 1 if and only if $\boldsymbol{z}_t^\top \boldsymbol{\theta}^* > 0$. Thus, for an algorithm π to be effective, its classification should be close to this linear one. To quantify the performance of π and compare it with the optimal classifier, we first define *π-induced estimator in period t*, $\boldsymbol{\theta}_t^\pi$, to be the maximizer of

$$P\left(a_t^\pi = 1, \boldsymbol{z}_t^\top \boldsymbol{\theta} > 0 \middle| \mathcal{H}_t\right) + P\left(a_t^\pi = 2, \boldsymbol{z}_t^\top \boldsymbol{\theta} \leq 0 \middle| \mathcal{H}_t\right), \tag{1}$$

i.e., the best linear approximation of π as a classifier. We will show that the regret of π can be bounded from below by the estimation error of $\boldsymbol{\theta}_t^\pi$ as an estimator of $\boldsymbol{\theta}^*$. The following lemma provides a lower bound for $R_t^\pi(\boldsymbol{\theta}^*, F)$, and hence $\sup_{\boldsymbol{\theta}^* \in \boldsymbol{\Theta}^*} R_t^\pi(\boldsymbol{\theta}^*, F)$.

Lemma 3.1. *For any $(\boldsymbol{\theta}^*, F) \in (\boldsymbol{\Theta}^*, \mathcal{F}_2)$, $h \geq 0$, $\mathcal{H}_t$ and maximizer $\boldsymbol{\theta}_t^\pi$,*

$$\begin{aligned} R_t^\pi(\boldsymbol{\theta}^*, F) \geq & \frac{h}{2} P\left(\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}^*\right)\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi\right) < 0\right) \\ & - hP\left(\left|\boldsymbol{z}_t^\top \boldsymbol{\theta}^*\right| < h,\ \operatorname{sgn}\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}^*\right) \neq \operatorname{sgn}\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi\right)\right), \end{aligned}$$

where $\operatorname{sgn}(\cdot)$ is the sign of a real number with $\operatorname{sgn}(0) = 0$.

We will apply Lemma 3.1 to obtain insightful lower bounds under various choices of F in Sections 3.1.1 and 3.1.2. Since any vector of the same direction as $\boldsymbol{\theta}_t^\pi$ would also be a maximizer of 1, we will only care about how well the direction of $\boldsymbol{\theta}_t^\pi$ estimates that of $\boldsymbol{\theta}^*$ and primarily discuss *estimation error* of the form

$$L_{\boldsymbol{\Sigma}}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_t^\pi) := E\left[\left\|\frac{\boldsymbol{\theta}^*}{\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}}} - \frac{\boldsymbol{\theta}_t^\pi}{\|\boldsymbol{\theta}_t^\pi\|_{\boldsymbol{\Sigma}}}\right\|_{\boldsymbol{\Sigma}}^2\right],$$

where $\boldsymbol{\Sigma}$ is positive definite and $\|\boldsymbol{\theta}\|_{\boldsymbol{\Sigma}} = \sqrt{\boldsymbol{\theta}^\top \boldsymbol{\Sigma} \boldsymbol{\theta}}$. We will drop the subscript of L when $\boldsymbol{\Sigma}$ is an identity matrix. Let $\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ denote a centered normal distribution with covariance matrix $\boldsymbol{\Sigma}$ and $\boldsymbol{I}_n$ represent the n-dimensional identity matrix. We next derive detailed bounds for any F such that $\boldsymbol{z}_t$ is (a) a normal or truncated normal random vector in Section 3.1.1, and (b) a mixture of normal or truncated normal random vectors in Section 3.1.2.

3.1.1. When $\boldsymbol{z}_t$ is (Truncated) Normal

Theorem 3.2. *Suppose that $\boldsymbol{z}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma}$ is positive-definite. $E[R_t^\pi(\boldsymbol{\theta}^*, F)]$ dominates the estimation error as*

$$E[R_t^\pi(\boldsymbol{\theta}^*, F)] = \Omega\left(\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}} \cdot L_{\boldsymbol{\Sigma}}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_t^\pi)\right).$$

Endowing $\boldsymbol{\Theta}^*$ with any prior ν, we have

$$R_{[T]}^\pi(\boldsymbol{\Theta}^*, \{F\}) = \Omega\left(\sum_{t=1}^T \inf_{\boldsymbol{\theta}_t \in \hat{\boldsymbol{\Theta}}_t} E_\nu\left[\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}} \cdot L_{\boldsymbol{\Sigma}}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_t)\right]\right),$$

for any set of estimators $\hat{\boldsymbol{\Theta}}_t$ that includes $\{\boldsymbol{\theta}_t^\pi : \pi \in \Pi\}$ and distribution ν on $\boldsymbol{\Theta}^*$.

The requirement of $\boldsymbol{z}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ can be easily satisfied by many joint distributions of $(\boldsymbol{x}_{t1}, \boldsymbol{x}_{t2})$, which may require neither $\boldsymbol{x}_{t1}$ nor $\boldsymbol{x}_{t2}$ to be normal. The positive-definiteness of $\boldsymbol{\Sigma}$ on $\boldsymbol{\Theta}^*$ guarantees the denominators in the theorem are non-zero. When $\boldsymbol{\Sigma}$ is positive-semidefinite, we may apply Theorem 3.2 with $\boldsymbol{\Sigma} + \varepsilon \boldsymbol{I}_d$, which is positive-definite, and let $\varepsilon \to 0$.

Theorem 3.2 establishes a lower bound for the single-period regret for a given $(\boldsymbol{\theta}^*, F)$ and algorithm π. Indeed, $L_{\boldsymbol{\Sigma}}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_t^\pi)$ captures whether the direction of $\boldsymbol{\theta}^*$ is correctly identified, which suffices to indicate the optimal arm, while the magnitude of $\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}}$ scales up the magnitude of regret if the sub-optimal arm is selected. A trade-off sets in here as with a larger $\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}}$, the direction of $\boldsymbol{\theta}^*$ can be learnt to a higher level of accuracy (i.e., smaller $L_{\boldsymbol{\Sigma}}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_t^\pi)$), but choosing the sub-optimal arm would also incur a higher regret.

We allow $\hat{\boldsymbol{\Theta}}_t$ to be any superset of $\{\boldsymbol{\theta}_t^\pi : \pi \in \Pi\}$ to facilitate the evaluation of the bound, and the smaller the $\hat{\boldsymbol{\Theta}}_t$ is, the tighter the lower bound will be. For example, $\boldsymbol{\theta}_t^\pi$ is a function of $\mathcal{H}_t$, but we may allow estimators in $\hat{\boldsymbol{\Theta}}_t$ to make use of a larger information set, e.g., information on the realized rewards from unchosen arms as well, i.e., functions that map $\{(\boldsymbol{x}_{sa}, y_{sa}) : 1 \le s \le t-1, a = 1, 2\} \supseteq \mathcal{H}_t$ to a vector in $\mathbb{R}^d$.

A special case. For the special case where $\boldsymbol{\Sigma} = \frac{1}{d}\boldsymbol{I}_d$ and the support of ν is on a sphere with a radius $r > 0$, we can simplify the bound in Theorem 3.2 as $\Omega\left(\frac{1}{r\sqrt{d}} \cdot \sum_{t=1}^{T} \inf_{\boldsymbol{\theta}_t \in \hat{\boldsymbol{\Theta}}_t} E_\nu\left[\|\boldsymbol{\theta}^* - \boldsymbol{\theta}_t\|_2^2\right]\right)$. The Bayesian 2-norm risk $E_\nu\left[\|\boldsymbol{\theta}^* - \boldsymbol{\theta}_t\|_2^2\right]$ is well understood for various classes of estimation problems such as linear and sparse linear regressions. Thus, it allows us to apply estimation theory directly to calculate lower bounds for stochastic linear bandit problems. For instance, Example 13.1 in (Duchi, 2016) establishes that, for any $r > 0$, there exists a distribution ν supported on a sphere of radius r such that, for $d \gg 1$, $\inf_{\boldsymbol{\theta}_t \in \hat{\boldsymbol{\Theta}}_t} E_\nu\left[\|\boldsymbol{\theta}^* - \boldsymbol{\theta}_t\|_2^2\right]$ is of order $\Omega\left(r^2\left\{\frac{1}{2} - \frac{16r^2}{d}E\left[\lambda_{\max}\left(\sum_{s=1}^{t}\sum_{a=1}^{2}\boldsymbol{x}_{sa}\boldsymbol{x}_{sa}^\top\right)\right]\right\}\right)$, where $\hat{\boldsymbol{\Theta}}_t$ is the set of all functions that map $\{(\boldsymbol{x}_{sa}, y_{sa}) : 1 \le s < t, a = 1, 2\} \supseteq \mathcal{H}_t$ to a vector in $\mathbb{R}^d$ and $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue of a matrix. For $(\boldsymbol{x}_{t1}, \boldsymbol{x}_{t2}) \sim \mathcal{N}\left(\boldsymbol{0}, \frac{1}{2d}\boldsymbol{I}_{2d}\right)$,

$$E\left[\lambda_{\max}\left(\sum_{s=1}^{T}\sum_{a=1}^{2}\boldsymbol{x}_{sa}\boldsymbol{x}_{sa}^\top\right)\right] \le \frac{2t}{d},$$

and our bound becomes

$$R_{[T]}\left(\boldsymbol{\Theta}^*, \{F\}\right) = \Omega\left(\frac{r}{\sqrt{d}}\sum_{t=1}^{T}\left(\frac{1}{2} - \frac{32r^2t}{d^2}\right)\right),$$

which is $\Omega\left(\max\left\{\sqrt{dT}, d^{3/2}\right\}\right)$ at $r = \min\left\{\frac{d}{16\sqrt{T}}, 1\right\}$.

Truncated Normal. Note that under any F such that $\boldsymbol{z}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$, $\boldsymbol{z}_t$ must have an unbounded support, while in most literature, $\mathcal{F}_2$ is assumed to only include distributions such that $\boldsymbol{x}_{ta}$ are bounded. Proposition 3.3 establishes that the bounds in Theorem 3.2 still apply when $\boldsymbol{z}_t$ is a truncated normal random vector so that F can have a bounded support. Indeed, a normal distribution behaves very much like a bounded one due to its thin tails.

Proposition 3.3. *Suppose that $\boldsymbol{z}_t$ is a $\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ random vector truncated in an ℓ^p-ball of radius $M > 0$ centered at $\boldsymbol{0}$, where $\boldsymbol{\Sigma}$ is positive-definite and $\mathbb{E}\left(\|\boldsymbol{x}\|_2\right) \le \frac{M}{2}$, then the bounds in Theorem 3.2 apply.*

3.1.2. When $\boldsymbol{z}_t$ is a Mixture of (Truncated) Normal

The bounds derived in the previous section is easy to evaluate and thus useful when ν is supported on a high-dimensional sphere, i.e., when $\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}}$ is a constant. An even simpler form of the lower bound where we can still have 2-norm error terms is to let $\boldsymbol{z}_t$ be a mixture of 2-dimensional normal distributions and ν supported on products of circles.

We partition the elements in $\boldsymbol{\theta}$ into $\lfloor d/2 \rfloor$ number of 2-dimensional vectors as $\boldsymbol{\theta} = \left(\boldsymbol{\theta}^1, \boldsymbol{\theta}^2, \dots, \boldsymbol{\theta}^{\lceil d/2 \rceil}\right)$, where $\boldsymbol{\theta}^i = (\theta_{2i-1}, \theta_{2i})$, $1 \le i \le \left\lfloor \frac{d}{2} \right\rfloor$, and $\boldsymbol{\theta}^{\lceil d/2 \rceil} = \theta_d$ if d is an odd number. Similarly, we can partition $\boldsymbol{z}_t = \left(\boldsymbol{z}_t^1, \boldsymbol{z}_t^2, \dots, \boldsymbol{z}_t^{\lceil d/2 \rceil}\right)$. Suppose that $z_{t,d} = 0$ if d is an odd number and $\boldsymbol{z}_t$ can only take values of the form $\left(\boldsymbol{0}, \boldsymbol{0}, \cdots, \boldsymbol{z}_t^i, \cdots, \boldsymbol{0}\right)$, $1 \le i \le \lfloor d/2 \rfloor$, i.e., only one of the $\boldsymbol{\theta}^{*1}, \dots, \boldsymbol{\theta}^{*\lfloor d/2 \rfloor}$ can be learnt in each period. Indeed, the problem can be regarded as a collection of $d/2$ number of 2-dimensional ones, each being played for a probabilistic number of times.

Theorem 3.4. *Suppose that $\boldsymbol{z}_t = \left(\boldsymbol{0}, \boldsymbol{0}, \cdots, \boldsymbol{z}_t^i, \cdots, \boldsymbol{0}\right)$, $1 \le i \le \left\lfloor \frac{d}{2} \right\rfloor$, with equal probability and $\boldsymbol{z}_t^i \sim \mathcal{N}\left(\boldsymbol{0}, \sigma^2\boldsymbol{I}_2\right)$. $E[R_t^\pi(\boldsymbol{\theta}^*, F)]$ dominates the estimation error as*

$$E[R_t^\pi(\boldsymbol{\theta}^*, F)] = \Omega\left(\frac{\sigma}{d}\sum_{i=1}^{\lfloor d/2 \rfloor}\left\|\boldsymbol{\theta}^{*i}\right\|_2 \cdot L(\boldsymbol{\theta}^{*i}, (\boldsymbol{\theta}_t^\pi)^i)\right).$$

Endowing $\boldsymbol{\Theta}^*$ with any prior ν, we have

$$\begin{aligned} &R_{[T]}^\pi\left(\boldsymbol{\Theta}^*, \{F\}\right) \\ &= \Omega\left(\frac{\sigma}{d}\sum_{t=1}^{T}\inf_{\boldsymbol{\theta}_t \in \hat{\boldsymbol{\Theta}}_t}\sum_{i=1}^{\lfloor d/2 \rfloor}E_\nu\left[\left\|\boldsymbol{\theta}^{*i}\right\|_2 \cdot L(\boldsymbol{\theta}^{*i}, (\boldsymbol{\theta}_t)^i)\right]\right), \end{aligned}$$

for any set of estimators $\hat{\boldsymbol{\Theta}}_t$ that includes $\{\boldsymbol{\theta}_t^\pi : \pi \in \Pi\}$ and distribution ν on $\boldsymbol{\Theta}^*$.

For the special case where the support of ν is in a product of circles

$$\left\{\boldsymbol{\theta}^* : \begin{array}{c}\left\|\boldsymbol{\theta}^{*i}\right\|_2 = r,\ 1 \le i \le \left\lfloor \frac{d}{2} \right\rfloor, \\ \theta_d = 0 \text{ if } d \text{ is an odd number}\end{array}\right\}$$

for some $r > 0$, we have that

$$R_{[T]}^\pi\left(\boldsymbol{\Theta}^*, \{F\}\right) = \Omega\left(\frac{\sigma}{rd} \cdot \sum_{t=1}^{T}\inf_{\boldsymbol{\theta} \in \hat{\boldsymbol{\Theta}}_t}\left\{E_\nu\left[\|\boldsymbol{\theta}^* - \boldsymbol{\theta}\|_2^2\right]\right\}\right),$$

which also takes the form of an ℓ_2-risk.

We also have a truncated version.

Proposition 3.5. *If $\boldsymbol{z}_t = \left(\boldsymbol{0}, \boldsymbol{0}, \cdots, \boldsymbol{z}_t^i, \cdots, \boldsymbol{0}\right)$, $1 \le i \le \left\lfloor \frac{d}{2} \right\rfloor$, with equal probability and $\boldsymbol{z}_t^i$ is a $\mathcal{N}(\boldsymbol{0}, \sigma^2\boldsymbol{I}_2)$ random* *vector truncated in an ℓ^p-ball centered at $\mathbf{0}$ with radius M such that $E\left(\left\|\boldsymbol{z}_t^i\right\|_2\right) < \frac{M}{2}$, then the bounds in Theorem 3.4 apply.*

3.2. The K-Armed Problems

We first note that all the lower bounds derived in the previous section apply directly to their corresponding K-armed problems if we can select $F \in \mathcal{F}_K$ such that the arms can be divided into two groups of arms, where each group has the same expected reward.

To characterize how much harder a general K-armed problem is than a 2-armed problem, we first describe how we can encode a collection of N number of 2-armed problems into a single 2^N-armed problem following (Li et al., 2019). Note that there are 2^N distinct combinations of actions in the N problems. Suppose that the nth 2-armed problem has a true parameter $\boldsymbol{\theta}^*(n)$ and the context is $(\boldsymbol{x}_{t1}(n), \boldsymbol{x}_{t2}(n))$ in period t. Consider a 2^N-armed problem with $\boldsymbol{\theta}^* = (\boldsymbol{\theta}^*(n), \ldots, \boldsymbol{\theta}^*(N))$ and $\boldsymbol{x}_{ta} = \left(\boldsymbol{x}_{tb_1(a)}(1), \ldots, \boldsymbol{x}_{tb_N(a)}(N)\right)$, where $b_n(a) \in \{1, 2\}$, $n \geq 1$, is uniquely determined by the binary representation $a = 1 + \sum_{i=1}^{\infty}(b_i(a) - 1) \cdot 2^{i-1}$, $1 \leq a \leq 2^N$. Then the regret of selecting arm a in the 2^N-armed problem is the same as the sum of the regrets of the collection of 2-armed problems if $b_n(a)$ is selected from the nth problem, $1 \leq n \leq N$. Figure 1 demonstrates an example with $N = 3$.

In general, as K is not necessarily a power of 2, let $N = \lfloor \log_2(K) \rfloor$,

$$\boldsymbol{\theta}^* = (\boldsymbol{\theta}^*(1), \cdots, \boldsymbol{\theta}^*(N), \mathbf{0}), \tag{2}$$

$$\boldsymbol{x}_{ta} = \left(\boldsymbol{x}_{tb_1(a)}(1), \cdots, \boldsymbol{x}_{tb_N(a)}(N), \mathbf{0}\right), \tag{3}$$

where $\boldsymbol{x}_{tb_n(a)}(n)$ and $\boldsymbol{\theta}^*(n) \in \mathbb{R}^{\lfloor d/N \rfloor}$ and $\mathbf{0}$ is of dimension $d - N \cdot \lfloor d/N \rfloor$. Then,

$$R^\pi(\boldsymbol{\theta}^*, F) = \sum_{n=1}^{\lfloor \log_2(K) \rfloor} \max\left\{\boldsymbol{x}_{t1}(n)^\top \boldsymbol{\theta}^*(n), \boldsymbol{x}_{t2}(n)^\top \boldsymbol{\theta}^*(n)\right\} - \boldsymbol{x}_{tb_n(a_t^\pi)}^\top(n)\boldsymbol{\theta}^*(n).$$

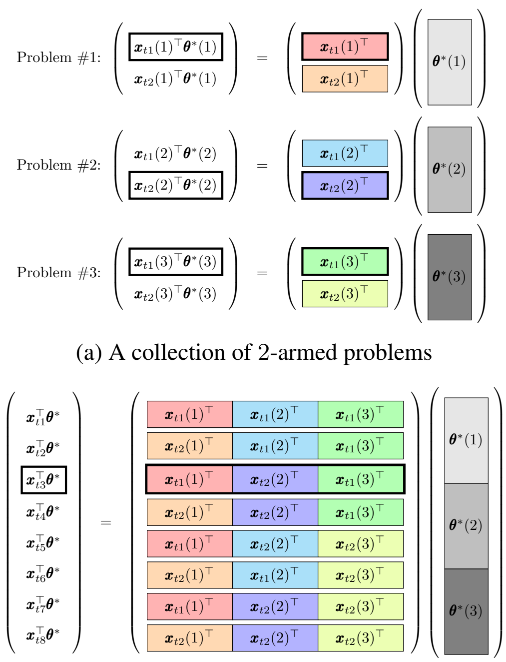
Let $\boldsymbol{z}_t(n) = \boldsymbol{x}_{t1}(n) - \boldsymbol{x}_{t2}(n)$ and $\boldsymbol{\theta}(n) = \boldsymbol{\theta}_t^\pi(n)$ be a maximizer of

$$\sum_{a=1}^{2} P\left(b_n(a_t^\pi) = a, (-1)^a \boldsymbol{z}_t(n)^\top \boldsymbol{\theta}(n) > 0\right),$$

for each $1 \leq n \leq N$. Theorem 3.6 generalizes Theorem 3.2.

Theorem 3.6. *Suppose that $\boldsymbol{z}_t(n) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma}$ is positive-definite and (2)–(3) holds.*

$$E[R_t^\pi(\boldsymbol{\theta}^*, F)] = \Omega\left(\sum_{n=1}^{N} \|\boldsymbol{\theta}^*(n)\|_{\boldsymbol{\Sigma}} \cdot L_{\boldsymbol{\Sigma}}(\boldsymbol{\theta}^*(n), \boldsymbol{\theta}_t^\pi(n))\right).$$



(a) A collection of 2-armed problems

(b) An 8-armed problem

Figure 1. Selecting arms $(a_{1t}, a_{2t}, a_{3t}) = (1, 2, 1)$ in (a) yields the same regret as selecting arm 3 in (b). Arm selection combinations in (a) are in one-to-one correspondence with the arms in (b).

Endowing $\boldsymbol{\Theta}^*$ with any prior ν,

$$R_{[T]}(\boldsymbol{\Theta}^*, \{F\}) = \Omega\left(\sum_{t=1}^{T}\sum_{n=1}^{N} \inf_{\boldsymbol{\theta}_t(n) \in \hat{\boldsymbol{\Theta}}_t(n)} E_\nu\left[\|\boldsymbol{\theta}^*(n)\|_{\boldsymbol{\Sigma}} \cdot L_{\boldsymbol{\Sigma}}(\boldsymbol{\theta}^*(n), \boldsymbol{\theta}_t(n))\right]\right),$$

for any set of estimators $\hat{\boldsymbol{\Theta}}_t(n)$ that includes $\{\boldsymbol{\theta}_t^\pi(n) : \pi \in \Pi\}$, $1 \leq i \leq N$.

When $\boldsymbol{z}_{nt}$ follows the distribution in Proposition 3.3, Theorem 3.4, or Proposition 3.5, similar generalizations can be obtained. We omit their statements and only present Proposition 3.7 below as a representative, which is a generalization of the special case discussed in Section 3.1.1. As one can see, there is an extra factor of $\sqrt{\log_2(K)}$.

Proposition 3.7. *When F satisfies conditions in Theorem 3.6 at $\boldsymbol{\Sigma} = \frac{1}{d}\boldsymbol{I}_{\lfloor d/\log_2(K) \rfloor}$ and $d \gg 1$,*

$$R_{[T]}(\boldsymbol{\Theta}^*, \{F\}) = \Omega\left(\max\left\{\sqrt{dT\log_2(K)}, d^{3/2}\right\}\right).$$

3.3. Comparing With an Upper Bound

Suppose that F satisfies the conditions in Theorem 3.6. Consider a generic greedy algorithm that based on a single estimator $\hat{\boldsymbol{\theta}}_t = (\hat{\boldsymbol{\theta}}_t(1), \ldots, \hat{\boldsymbol{\theta}}_t(N), \mathbf{0})$ in period t, i.e., one selects an arm with the highest expected reward pretending the true parameter is $\hat{\boldsymbol{\theta}}_t$, i.e., $a_t = \arg\max_{1 \leq a \leq K}\left\{\boldsymbol{x}_{ta}^\top \hat{\boldsymbol{\theta}}_t\right\}$.

Then, the expected regret in period t under this greedy algorithm can be calculated as

$$\begin{aligned}&E[R_t^\pi(\boldsymbol{\theta}^*,F)]\\=&\sum_{n=1}^N E\left[|\boldsymbol{z}_t(n)^\top\boldsymbol{\theta}^*(n)|1_{\{(\boldsymbol{z}_t(n)^\top\boldsymbol{\theta}^*(n))(\boldsymbol{z}_t(n)^\top\boldsymbol{\theta}^*(n))<0\}}\right]\\=&O\left(\sum_{n=1}^N\|\boldsymbol{\theta}^*(n)\|_{\boldsymbol{\Sigma}}\cdot L_{\boldsymbol{\Sigma}}(\boldsymbol{\theta}^*(n),\hat{\boldsymbol{\theta}}_t(n))\right),\end{aligned}$$

and hence,

$$\begin{aligned}&R_{[T]}^\pi(\boldsymbol{\Theta}^*,\{F\})\\=&O\left(\sum_{t=1}^T\sup_{\boldsymbol{\theta}^*\in\boldsymbol{\Theta}^*}\sum_{n=1}^N\|\boldsymbol{\theta}^*(n)\|_{\boldsymbol{\Sigma}}\cdot L_{\boldsymbol{\Sigma}}(\boldsymbol{\theta}^*(n),\hat{\boldsymbol{\theta}}_t(n))\right).\end{aligned}$$

Note that the upper bounds share the same form as the lower bounds in Theorem 3.6. Thus, if the worst-case risk of the $\hat{\boldsymbol{\theta}}_t$ is of the same order as the Bayesian risk of the Bayesian estimator, our lower bound is tight.

4. Applications of the Generic Bound

4.1. The Basic Model

4.1.1. When $K=2$

For 2-armed problems, we specify our choice of $(F,\nu,\hat{\boldsymbol{\Theta}}_t)$ as follows.

- $F\in\mathcal{F}_2$: Draw $\boldsymbol{z}_t$ from the mixture of truncated normal random vectors as described in Proposition 3.5 at $p=2$ and assign $(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2})=(\mathbf{0},-\boldsymbol{z}_t)$ or $(\boldsymbol{z}_t,\mathbf{0})$ with equal probability.
- ν: A uniform distribution on a product of circles $\boldsymbol{\Theta}_r^*=\{\boldsymbol{\theta}^*=\left(\boldsymbol{\theta}^{*1},\ldots,\boldsymbol{\theta}^{*\lceil d/2\rceil}\right):\boldsymbol{\theta}^{*i}\in\mathbb{R}^2,\ \left\|\boldsymbol{\theta}^{*i}\right\|_2=r,\ 1\le i\le\left\lfloor\frac{d}{2}\right\rfloor,\theta_d=0\text{ if }d\text{ is an odd number}\}$ for some $r\ge 0$ specified later.
- $\hat{\boldsymbol{\Theta}}_t$: The set of all functions that map

  $$\overline{\mathcal{H}}_t=\{(\boldsymbol{x}_{sa},y_{sa}):a=1,2,\ \ 1\le s\le t-1\}\supseteq\mathcal{H}_t$$

  to a vector in $\mathbb{R}^d$. It is easy to verify that that $\hat{\boldsymbol{\Theta}}_t$ meets the requirements in Theorem 3.4.

Proposition 4.1. $R_{[T]}(\boldsymbol{\Theta}^*,\{F\})=\sum_{t=1}^T\Omega\left(\sigma\min\left\{r,\frac{d}{rt}\right\}\right)$

Indeed, Proposition 4.1 follows a simple procedure of calculation under our framework. This recovers some well-known bounds in the literature.

- Auer (2002) and Chu et al. (2011) studied adversarial contexts with $\|\boldsymbol{x}_t\|_2\le 1$ and thus includes stochastic contexts and our choice of F at $\sigma=\frac{1}{\sqrt{2}}$ and $M=1$, and $\boldsymbol{\Theta}^*=\{\boldsymbol{\theta}^*:\|\boldsymbol{\theta}^*\|_2\le 1\}$ so the support of our ν is indeed a subset of $\boldsymbol{\Theta}^*$ for any $r\le\frac{1}{\sqrt{d}}$. At $r=\sqrt{\frac{d}{T}}$ (assuming $d<\sqrt{T}$), our lower bound $\sum_{t=1}^T\Omega\left(\min\left\{r,\frac{d}{rt}\right\}\right)=\Omega\left(\sqrt{dT}\right)$ which coincides with theirs.
- Goldenshluger & Zeevi (2013) further imposes a margin condition $P\left(\left|(\boldsymbol{x}_{t1}-\boldsymbol{x}_{t2})^\top\boldsymbol{\theta}^*\right|\le\rho\right)\le L\rho$ for some constant L independent of (T,d) and any $\rho>0$ on the set of distributions. At $r=1$, our choice of $(F,\nu,\hat{\boldsymbol{\Theta}}_t)$ satisfies the requirements and our lower bound $\Omega\left(d+d\log(T/d)\right)$ is slightly tighter than theirs $\Omega(\log(T))$ and the same as $\Omega(d\log(T))$ as suggested in Oh et al. (2021).

4.1.2. When $K\ge 2$

Our choice of $(F,\nu,\hat{\boldsymbol{\Theta}}_t)$ is as follows.

- F: Draw $\boldsymbol{z}_t(n)=(\boldsymbol{z}_t^1(n),\ldots,\boldsymbol{z}_t^{\lceil d/2\rceil}(n)),1\le n\le N=\lfloor\log_2(K)\rfloor$, from the distribution described in Proposition 3.3 with $p=2$, assign $(\boldsymbol{x}_{t1}(n),\boldsymbol{x}_{t2}(n))=(\mathbf{0},-\boldsymbol{z}_t(n))$ or $(\boldsymbol{z}_t(n),\mathbf{0})$ with equal probability, and generate $\boldsymbol{x}_{t1},\ldots,\boldsymbol{x}_{tK}$, following (2)–(3).
- ν: A uniform distribution on $\boldsymbol{\Theta}_r^*(1)\times\cdots\times\boldsymbol{\Theta}_r^*(N)\times\{\mathbf{0}\}$ where $\boldsymbol{\Theta}_r^*(n)$ is a product of circles $\{\boldsymbol{\theta}^*(n)=\left(\boldsymbol{\theta}^{*1}(n),\ldots,\boldsymbol{\theta}^{*\lceil d_0/2\rceil}(n)\right):\boldsymbol{\theta}^{*i}(n)\in\mathbb{R}^2,\ \left\|\boldsymbol{\theta}^{*i}(n)\right\|_2=r,\ 1\le i\le\left\lfloor\frac{d_0}{2}\right\rfloor,\theta_{d_0}=0\text{ if }d_0\text{ is an odd number}\}$, with $d_0=\lfloor d/N\rfloor$.
- $\hat{\boldsymbol{\Theta}}_t(n)$: The set of all functions that map $\overline{\mathcal{H}}_t\cup\{\boldsymbol{\theta}^*(n'):1\le i'\ne i\le N\}$ to a vector in $\mathbb{R}^{d_0}$.

Following a similar argument as that for the 2-armed case and letting $\hat{\boldsymbol{\theta}}_t(n)$ be the minimizer of $\inf_{\boldsymbol{\theta}_t(n)\in\hat{\boldsymbol{\Theta}}_t(n)}\sum_{i=1}^{\lfloor d_0/2\rfloor}E_\nu\left[\left\|\boldsymbol{\theta}^{*i}(n)-\boldsymbol{\theta}_t^i(n)\right\|_2\right]$, we have

$$E\left[\left\|\boldsymbol{\theta}^{i*}(n)-\boldsymbol{\theta}^i(n)\right\|_2^2\right]=\Omega\left(\min\left\{r^2,\frac{\lfloor d/\log_2(K)\rfloor}{t}\right\}\right).$$

Hence,

$$R_{[T]}(\boldsymbol{\Theta}^*,\{F\})=\sum_{t=1}^T\Omega\left(\min\left\{r\log_2(K),\frac{d}{rt}\right\}\right),$$

which is $\Omega\left(\sqrt{dT\log(K)}\right)$ at $r=\sqrt{\frac{d}{T\log_2(K)}}$ and is $\Omega\left(d\log\left(\frac{T\log_2(K)}{d}\right)\right)$ at $r=1$.

4.1.3. With a Batch Constraint

Here we revisit the setting in (Han et al., 2020) where feedbacks can be received in M batches and the decision maker needs to decide the *grids*, i.e., the M periods, $t_1 < t_2 < \cdots < t_M = T$, in which feedbacks are revealed, before pulling any arm.

Proposition 4.2.

$$R_{[T]}(\boldsymbol{\Theta}^*, \{F\}) = \max_{m=1,\dots,M} \Omega\left(t_m \min\left\{r, \frac{d}{rt_{m-1}}\right\}\right)$$

The proof of the proposition is almost identical to Proposition 4.1 and hence omitted. Maximizing over r we can recover the lower bound for stochastic linear contextual bandit in their work.

4.2. The Sparse Linear Bandit Problem

Abundant work on sparse linear bandit problems have focused on constructing an algorithm and establishing its effectiveness by providing a regret upper bound for it. To name a few, Oh et al. (2021), Wang et al. (2020), and Li et al. (2021) achieve the best upper bounds of order $\tilde{O}(\sqrt{s_0 T})$, where $s_0 = \|\boldsymbol{\theta}^*\|_0, \forall \boldsymbol{\theta}^* \in \boldsymbol{\Theta}^*$, is the sparse parameter. Since the conditions on $(F, \nu, \boldsymbol{\Theta}^*)$ in Wang et al. (2020) are easy to meet and compare, we will compare our lower bound with their upper bound. When $K = 2$, we choose $(F, \nu, \hat{\boldsymbol{\Theta}}_t)$ as follows.

- F: $(\boldsymbol{x}_{t1}, \boldsymbol{x}_{t2}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_{2d})$.
- ν: A uniform distribution on
$$\boldsymbol{\Theta}_r^* = \{\boldsymbol{\theta}^* : \theta_i^* \in \{-r, 0, r\}, \|\boldsymbol{\theta}^*\|_0 = s_0\},$$
where s_0 is a sparse parameter and $r > 0$ will be specified later.
- $\hat{\boldsymbol{\Theta}}_t$: The same as in Section 4.1.

The lower bound in Theorem 3.2 can be simplified as $\Omega\left(\frac{1}{r\sqrt{s_0}} \sum_{t=1}^{T} \inf_{\boldsymbol{\theta}_t \in \hat{\boldsymbol{\Theta}}_t} \left\{E_\nu\left[\|\boldsymbol{\theta}^* - \boldsymbol{\theta}_t\|_2^2\right]\right\}\right)$.
Corollary 4 in (Duchi & Wainwright, 2013) establishes that, there exists a constant c independent of (T, d, s_0) such that

$$\inf_{\boldsymbol{\theta}_t \in \hat{\boldsymbol{\Theta}}_t} E_\nu\left[\|\boldsymbol{\theta}^* - \boldsymbol{\theta}_t\|_2^2\right] = \Omega\left(s_0 r^2 \left[1 - \frac{\frac{s_0 r^2}{d} E\left[\sum_{s=1}^{T}\sum_{a=1}^{2} \|\boldsymbol{x}_{sa}\|_2^2\right] + \log 2}{c s_0 \log(\frac{d}{s_0})}\right]\right).$$

Letting $r = \sqrt{\frac{c\log\left(\frac{d}{s_0}\right)}{8T}}$, we have $R_{[T]}(\boldsymbol{\Theta}^*, \{F\}) = \Omega\left(\sqrt{s_0 T \log\left(\frac{d}{s_0}\right)}\right)$, which only differs from the upper bound in (Wang et al., 2020) by a factor of $\log^3(T)$. For a general $K \geq 2$, a similar procedure yields

$$R_{[T]}(\boldsymbol{\Theta}^*, \{F\}) = \Omega\left(rT\sqrt{s_0 \log(K)}\left(1 - \frac{r^2 T}{c\log\left(\frac{d}{s_0}\right)}\right)\right),$$

which is $\Omega\left(\sqrt{s_0 T \log(K) \log(d/s_0)}\right)$ at $r = \sqrt{\frac{c\log(d/s_0)}{8T}}$.

4.3. Jointly Differentially Private Linear Bandit Problem

Consider a set of jointly differentially private linear bandit problems where $\mathcal{F}_K$ is the set of sub-Gaussian distributions with parameter 1, $\boldsymbol{\Theta}^* = \{\boldsymbol{\theta}^* : \|\boldsymbol{\theta}^*\|_2 \leq 1\}$, and Π is the set of (ε, δ)-jointly differentially private algorithms defined in Section 2. To calculate a lower bound for $K = 2$, we will choose the same F and ν as in Section 4.1 and $\hat{\boldsymbol{\Theta}}_t$ to be the set of all (ε, δ)-differentially private functions that map $\overline{\mathcal{H}}_t = \{(\boldsymbol{x}_{sa}, y_{sa}) : a = 1, 2, \ \ 1 \leq s \leq t-1\} \supseteq \mathcal{H}_t$ to a vector in $\mathbb{R}^d$.

Proposition 4.3. $R_{[T]}(\boldsymbol{\Theta}^*, \{F\}) = \Omega(\sum_{t=1}^{T} \min\{r, \frac{d^2}{rt^2(\varepsilon+\delta)^2}\})$.

Letting $r = \frac{d}{T(\varepsilon+\delta)}$, we obtain that $R_{[T]}(\boldsymbol{\Theta}^*, \{F\}) = \Omega\left(\frac{d}{\varepsilon+\delta}\right)$. Since the choice of $(F, \nu, \hat{\boldsymbol{\Theta}}_t)$ differs from that in Section 4.1 only in that $\hat{\boldsymbol{\Theta}}_t$ is restricted to a subset, a lower bound of $\Omega(\sqrt{dT})$ also applies. Thus, we obtain a lower bound of $\Omega\left(\sqrt{dT} + \frac{d}{\varepsilon+\delta}\right)$. An upper bound of $\tilde{O}\left(\sqrt{dT} + \frac{d\log(1/\delta)}{\varepsilon}\right)$ can be obtained from combining the estimator developed in (Cai et al., 2020) and the data batching scheme in (Han et al., 2020), which differs with our lower bounds only by factors of $\log(T)$ and $\log(\delta)$. For a general K, following a similar analysis as above and as in the previous sections, we have $R_{[T]}(\boldsymbol{\theta}^*, F) = \Omega\left(\sqrt{dT\log_2(K)} + \frac{d}{\varepsilon+\delta}\right)$.

4.4. Non-interactive Locally Differentially Private Sparse Linear Bandits

Consider a set of problems where $\mathcal{F}_K$ is the set of sub-Gaussian distributions with parameter 1,

$$\boldsymbol{\Theta}^* = \{\boldsymbol{\theta}^* : \|\boldsymbol{\theta}^*\|_2 \leq 1, \|\boldsymbol{\theta}^*\|_0 = 1\},$$

a set of sparse vectors, and Π is the set of non-interactive ε-locally differentially private algorithms.

When $K = 2$, to apply Theorem 3.2, we let $\boldsymbol{X}_t = (\boldsymbol{x}_{t1}, \boldsymbol{x}_{t2}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}_{2d})$, ν be a uniform distribution on a $2d$-point set

$$\begin{aligned}\boldsymbol{\Theta}_r^* &:= \{\boldsymbol{\theta}^* : \|\boldsymbol{\theta}^*\|_2 = r, \|\boldsymbol{\theta}^*\|_0 = 1\} \\ &= \{r\boldsymbol{e}_1, -r\boldsymbol{e}_1, \dots, r\boldsymbol{e}_d, -r\boldsymbol{e}_d\} \subseteq \boldsymbol{\Theta}^*\end{aligned}$$

for some $0 < r \leq 1$, where $\boldsymbol{e}_i$ is the vector with a 1 in the ith coordinate and 0's elsewhere, and $\hat{\boldsymbol{\Theta}}_t$ be the set of functions that map a non-interactive ε-LDP view of $\{(\boldsymbol{X}_s, y_{s1}, y_{s2}) : 1 \leq s \leq t-1\}$ to a vector in $\mathbb{R}^d$. The construction for a general K is similar.

Proposition 4.4. $R_{[T]}(\boldsymbol{\Theta}^*, \{F\}) = \Omega\left(r\log(K)\left[1 - \frac{T\log(K)r^2\varepsilon^2}{d(1-2r^2)\log(d/\log(K))} - \frac{\log(2)}{\log(d/\log(K))}\right]\right)$.

The bound reduces to $\Omega\left(\min\{\frac{\sqrt{d\log(d/\log(K))\log(K)T}}{\varepsilon}, T\}\right)$ at $r = \min\left\{\frac{\sqrt{d\log(d/\log(K)}}{2\varepsilon\sqrt{T\log(K)}}, 1\right\}$. From Sections 4.1 and 4.2, sparsity in $\boldsymbol{\Theta}^*$ reduces regrets in linear bandit problems from an order of $\sqrt{d}$ to an order of $\log(d)$. We show in this section that, in the presence of the ε-locally differentially privacy requirement, the regret is at least in the order of $\frac{\sqrt{d\log(d)T}}{\varepsilon}$, indicating that sparsity of $\boldsymbol{\Theta}^*$ may not reduce regrets in high-dimensional problems.

5. Conclusion

In this work, we establish the necessity of an accurate estimator in a low-regret algorithm for stochastic linear contextual bandit problems and demonstrate how our analysis leads to a reduction in studying lower bounds for bandit problems. Moreover, we have identified hard instances where contexts follow normal distributions (or their mixture) in various linear contextual bandit problems, suggesting developing algorithms against normal distributions might be a promising approach. Although our current instances are constructed with context following (truncated) normal distribution, all the methods can be readily applied to context with an ellipsoidal contoured context that does not highly concentrated around $\mathbf{0}$. In general, the gap between learning an optimal action and learning the entire reward function is interesting and deserves further investigation, and our work may serve as a first attempt on a systematic approach.

Acknowledgements

This work was supported in part by Hong Kong Research Grant Council (HKRGC) Grant 16214121, 16208120, 16200617, 16200019, 16200821. The authors thank the anonymous reviewers for their useful comments.

References

Acharya, J., Sun, Z., and Zhang, H. Differentially private assouad, fano, and le cam. In *Algorithmic Learning Theory*, pp. 48–78. PMLR, 2021.

Auer, P. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3(Nov):397–422, 2002.

Bastani, H. and Bayati, M. Online decision making with high-dimensional covariates. *Operations Research*, 68 (1):276–294, 2020.

Bastani, H., Bayati, M., and Khosravi, K. Mostly exploration-free algorithms for contextual bandits. *Management Science*, 67(3):1329–1349, 2021.

Cai, T. T., Wang, Y., and Zhang, L. The cost of privacy in generalized linear models: Algorithms and minimax lower bounds. *arXiv preprint arXiv:2011.03900*, 2020.

Chu, W., Li, L., Reyzin, L., and Schapire, R. Contextual bandits with linear payoff functions. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pp. 208–214. JMLR Workshop and Conference Proceedings, 2011.

Devroye, L., Mehrabian, A., and Reddad, T. The total variation distance between high-dimensional gaussians. *arXiv preprint arXiv:1810.08693*, 2018.

Duchi, J. Lecture notes for statistics 311/electrical engineering 377. *URL: https://stanford.edu/class/stats311/Lectures/full_notes. pdf. Last visited on*, 2:23, 2016.

Duchi, J. C. and Wainwright, M. J. Distance-based and continuum fano inequalities with applications to statistical estimation. *arXiv preprint arXiv:1311.2669*, 2013.

Duchi, J. C., Jordan, M. I., and Wainwright, M. J. Minimax optimal procedures for locally private estimation. *Journal of the American Statistical Association*, 113(521):182–201, 2018.

Goldenshluger, A. and Zeevi, A. A linear response bandit problem. *Stochastic Systems*, 3(1):230–261, 2013.

Han, Y., Zhou, Z., Zhou, Z., Blanchet, J., Glynn, P. W., and Ye, Y. Sequential batch learning in finite-action linear contextual bandits. *arXiv preprint arXiv:2004.06321*, 2020.

Han, Y., Liang, Z., Wang, Y., and Zhang, J. Generalized linear bandits with local differential privacy. *arXiv preprint arXiv:2106.03365*, 2021.

Kannan, S., Morgenstern, J., Roth, A., Waggoner, B., and Wu, Z. S. A smoothed analysis of the greedy algorithm for the linear contextual bandit problem. *arXiv preprint arXiv:1801.03423*, 2018.

Karwa, V. and Vadhan, S. Finite sample differentially private confidence intervals. *arXiv preprint arXiv:1711.03908*, 2017.

Kim, G.-S. and Paik, M. C. Doubly-robust lasso bandit. *arXiv preprint arXiv:1907.11362*, 2019.

Li, L., Chu, W., Langford, J., and Schapire, R. E. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World wide web*, pp. 661–670, 2010.

Li, W., Barik, A., and Honorio, J. A simple unified framework for high dimensional bandit problems. *arXiv preprint arXiv:2102.09626*, 2021.

Li, Y., Wang, Y., and Zhou, Y. Nearly minimax-optimal regret for linearly parameterized bandits. In *Conference on Learning Theory*, pp. 2173–2174. PMLR, 2019.

Oh, M.-h., Iyengar, G., and Zeevi, A. Sparsity-agnostic lasso bandit. In *International Conference on Machine Learning*, pp. 8271–8280. PMLR, 2021.

Ren, Z. and Zhou, Z. Dynamic batch learning in high-dimensional sparse linear contextual bandits. *arXiv preprint arXiv:2008.11918*, 2020.

Shariff, R. and Sheffet, O. Differentially private contextual linear bandits. *arXiv preprint arXiv:1810.00068*, 2018.

Wang, D. and Xu, J. On sparse linear regression in the local differential privacy model. In *International Conference on Machine Learning*, pp. 6628–6637. PMLR, 2019.

Wang, Y., Chen, Y., Fang, E. X., Wang, Z., and Li, R. Nearly dimension-independent sparse linear bandit over small action spaces via best subset selection. *arXiv preprint arXiv:2009.02003*, 2020.

Zheng, K., Cai, T., Huang, W., Li, Z., and Wang, L. Locally differentially private (contextual) bandits learning. *arXiv preprint arXiv:2006.00701*, 2020.

A. Appendix

Lemma A.1. *Let U and V be two standard normal random variables and $\beta = \arccos\left(\mathrm{Cov}(U,V)\right) > 0$. Then, for any $h > 0$ and $k \geq 1$,*

$$P\left(U > 0, -h < V < 0\right) < \frac{\beta}{2\pi}\left(1 - \frac{k-1}{k} e^{-\frac{h^2}{2\sin^2\left(\frac{1}{k}\left(\beta \wedge \frac{\pi}{2}\right)\right)}}\right).$$

Before a formal proof of the above lemma, we first provide some geometric intuition on it. The gray area in Figure 2 below describes the LHS of the inequality and is covered by the shaded cone plus the shaded circular sector, whose probabilities can be explicitly evaluated due to symmetric properties of 2-dimensional normal random vectors.

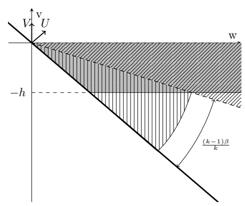


Figure 2.

Proof. Let $W = \frac{U - V\cos(\beta)}{\sin(\beta)}$. Then it is easy to show that $W \sim \mathcal{N}(0,1)$ and is independent of V. Furthermore, $\arctan\left(\frac{V}{W}\right)$ is uniform on $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, $P\left(\sqrt{V^2+W^2} < r\right) = 1 - e^{-\frac{r^2}{2}}$, and $\arctan\left(\frac{V}{W}\right)$ is independent of $\sqrt{V^2+W^2}$. For $\beta \leq \frac{\pi}{2}$,

$$\begin{aligned}
&P(U > 0, -h < V < 0) \\
&= P(\max\{-W\tan(\beta), -h\} < V < 0) \\
&\leq P\left(\max\{-W\tan(\beta), -h\} < V < -W\tan\left(\frac{\beta}{k}\right) < 0\right) + P\left(-W\tan\left(\frac{\beta}{k}\right) < V < 0\right) \\
&< P\left(-W\tan(\beta) < V < -W\tan\left(\frac{\beta}{k}\right), -h < V < 0, 0 < W < \frac{h}{\tan\left(\frac{\beta}{k}\right)}\right) + P\left(-W\tan\left(\frac{\beta}{k}\right) < V < 0\right) \\
&< P\left(-\beta < \arctan\left(\frac{V}{W}\right) < -\frac{\beta}{k}, \sqrt{V^2+W^2} < \frac{h}{\sin\left(\frac{\beta}{k}\right)}, V < 0\right) + P\left(\arctan\left(\frac{V}{W}\right) > -\frac{\beta}{k}, V < 0\right) \\
&= \frac{\beta}{2\pi}\left(1 - \frac{k-1}{k} e^{-\frac{h^2}{2\sin^2\left(\frac{\beta}{k}\right)}}\right).
\end{aligned}$$

Note that, at $\beta = \frac{\pi}{2}$, U and V are independent, and we obtain that the marginal distribution satisfies $P(-h < V < 0) < \frac{1}{2}\left(1 - \frac{k-1}{k} e^{-\frac{h^2}{2\sin^2\left(\frac{\pi}{2k}\right)}}\right)$. For $\beta > \frac{\pi}{2}$,

$$\begin{aligned}
P(U > 0, -h < V < 0) &= P(-h < V < \min\{-W\tan(\beta), 0\}) \\
&< P\left(W > \frac{h}{\tan(\beta)}\right) P\left(-h < V < 0\right) \\
&= P\left(W > \frac{h}{\tan(\beta)}\right) P\left(V < -h\right) \frac{P\left(-h < V < 0\right)}{1 - P\left(-h < V < 0\right)} \\
&< P\left(\frac{V}{W} < -\tan(\beta), V < 0\right) \frac{P\left(-h < V < 0\right)}{1 - P\left(-h < V < 0\right)} \\
&< \frac{\beta}{2\pi}\left(1 - \frac{k-1}{k} e^{-\frac{h^2}{2\sin^2\left(\frac{\pi}{2k}\right)}}\right).
\end{aligned}$$

□

The proof of Lemma 3.1 is basically exploiting the optimality of $\boldsymbol{\theta}_t^\pi$. In particular, $\boldsymbol{\theta}_t^\pi$ approximates π better than $\boldsymbol{\theta}^*$ does. As a consequence, when the difference of contexts, $\boldsymbol{z}_t$, falls in the cone $\left\{\boldsymbol{z} : \left(\boldsymbol{z}^\top \boldsymbol{\theta}_t^\pi\right)\left(\boldsymbol{z}^\top \boldsymbol{\theta}^*\right) < 0\right\}$ where $\boldsymbol{\theta}_t^\pi$ makes the wrong classification, π agrees with $\boldsymbol{\theta}_t^\pi$ more than $\boldsymbol{\theta}^*$. In other words, π makes more wrong classifications than correct ones in this cone and we can provide a regret lower bound based on this fact and the size of the cone. See a formal proof below.

Proof of Lemma 3.1. Note that

$$\begin{aligned}
-\boldsymbol{z}_t^\top \boldsymbol{\theta}^* 1_{\{a_t^\pi=1\}} 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le 0\}} &\ge -\boldsymbol{z}_t^\top \boldsymbol{\theta}^* 1_{\{a_t^\pi=1\}} 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi > 0,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le 0\}} \\
&\ge h 1_{\{a_t^\pi=1\}} 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi > 0,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le -h\}} \\
&= h 1_{\{a_t^\pi=1\}} 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi > 0,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le 0\}} - h 1_{\{a_t^\pi=1\}} 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi > 0,\ -h < \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le 0\}} \\
&\ge h 1_{\{a_t^\pi=1\}} 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi > 0,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le 0\}} - h 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi > 0,\ -h < \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le 0\}} \\
&= \frac{h}{2}\left(1 + 1_{\{a_t^\pi=1\}} - 1_{\{a_t^\pi=2\}}\right) 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi > 0,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le 0\}} - h 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi > 0,\ -h < \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le 0\}}.
\end{aligned}$$

Similarly, we have

$$\boldsymbol{z}_t^\top \boldsymbol{\theta}^* 1_{\{a_t^\pi=2\}} 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}^* > 0\}} \ge \frac{h}{2}\left(1 + 1_{\{a_t^\pi=2\}} - 1_{\{a_t^\pi=1\}}\right) 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi \le 0,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}^* > 0\}} - h 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi \le 0,\ 0 < \boldsymbol{z}_t^\top \boldsymbol{\theta}^* < h\}}.$$

Thus,

$$\begin{aligned}
R_t^\pi(\boldsymbol{\theta}^*, F) &= -\boldsymbol{z}_t^\top \boldsymbol{\theta}^* 1_{\{a_t^\pi=1\}} 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le 0\}} + \boldsymbol{z}_t^\top \boldsymbol{\theta}^* 1_{\{a_t^\pi=2\}} 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}^* > 0\}} \\
&\ge \frac{h}{2}\left[P\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi > 0,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le 0\right) + P\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi \le 0,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}^* > 0\right)\right] \\
&\quad -h\left[P\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi > 0, -h < \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le 0\right) + P\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi \le 0, 0 < \boldsymbol{z}_t^\top \boldsymbol{\theta}^* < h\right)\right] \\
&\quad + \frac{h}{2} E\left[\left(1_{\{a_t^\pi=1\}} - 1_{\{a_t^\pi=2\}}\right)\left(1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi > 0,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le 0\}} - 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi \le 0,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}^* > 0\}}\right)\right] \\
&\ge \frac{h}{2} P\left(\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}^*\right)\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi\right) < 0\right) - hP\left(\left|\boldsymbol{z}_t^\top \boldsymbol{\theta}^*\right| < h, \operatorname{sgn}\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi\right) \ne \operatorname{sgn}\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}^*\right)\right) \\
&\quad + \frac{h}{2} E\left[\left(1_{\{a_t^\pi=1\}} - 1_{\{a_t^\pi=2\}}\right)\left(1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi > 0\}} - 1_{\{\boldsymbol{z}_t^\top \boldsymbol{\theta}^* > 0\}}\right)\right]
\end{aligned}$$

and the lemma follows as the last term is equal to $\frac{h}{2}\left[P\left(a_t^\pi = 1,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi > 0\right) + P\left(a_t^\pi = 2,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi \le 0\right) - P\left(a_t^\pi = 1,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}^* > 0\right) - P\left(a_t^\pi = 2,\ \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \le 0\right)\right]$, which is non-negative due to the optimality of $\boldsymbol{\theta}_t^\pi$. □

Theorem 3.2 is established by evaluating the lower bound in Lemma 3.1. Lemma A.1 is applied to establish a concise lower bound.

Proof of Theorem 3.2. By Lemma 3.1 and the symmetry of the normal distribution, we have

$$\begin{aligned}
E\left[R_t^\pi(\boldsymbol{\theta}^*, F) | \mathcal{H}_t\right] &\ge h\left[\frac{1}{2} P\left(\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}^*\right)\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi\right) < 0 \middle| \mathcal{H}_t\right) - P\left(\left|\boldsymbol{z}_t^\top \boldsymbol{\theta}^*\right| \le h,\ \operatorname{sgn}(\boldsymbol{z}_t^\top \boldsymbol{\theta}^*) \ne \operatorname{sgn}(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi) \middle| \mathcal{H}_t\right)\right] \\
&= h\left[P\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi < 0 < \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \middle| \mathcal{H}_t\right) - 2P\left(0 < \boldsymbol{z}_t^\top \boldsymbol{\theta}^* < h,\ \operatorname{sgn}(\boldsymbol{z}_t^\top \boldsymbol{\theta}^*) = -\operatorname{sgn}(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi) = 1 \middle| \mathcal{H}_t\right)\right] \\
&= h\left[P\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi < 0 < \boldsymbol{z}_t^\top \boldsymbol{\theta}^* \middle| \mathcal{H}_t\right) - 2P\left(\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi < 0 < \boldsymbol{z}_t^\top \boldsymbol{\theta}^* < h \middle| \mathcal{H}_t\right)\right] \\
&\ge h\left[\frac{\beta_t^\pi}{2\pi} - \frac{\beta_t^\pi}{\pi}\left(1 - \frac{k-1}{k} e^{-\frac{h^2}{2\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}}^2 \sin^2\left(\frac{1}{k}\left(\beta \wedge \frac{\pi}{2}\right)\right)}}\right)\right] \\
&\ge \frac{h\beta_t^\pi}{\pi}\left[\frac{k-1}{k} e^{-\frac{h^2}{2\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}}^2 \sin^2\left(\frac{1}{k}\left(\beta \wedge \frac{\pi}{2}\right)\right)}} - \frac{1}{2}\right],
\end{aligned}$$

for $k > 1$ and $h > 0$, where $\beta_t^\pi = \arccos\left(\frac{(\boldsymbol{\theta}_t^\pi)^\top \boldsymbol{\Sigma}\boldsymbol{\theta}^*}{\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}}}\right)$ and the last inequality follows from Lemma A.1 in the Appendix at $U = \frac{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi}{Var[\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi]} = \frac{\boldsymbol{z}_t^\top \boldsymbol{\theta}_t^\pi}{\|\boldsymbol{\theta}_t^\pi\|_{\boldsymbol{\Sigma}}}$ and $V = \frac{\boldsymbol{z}_t^\top \boldsymbol{\theta}^*}{Var[\boldsymbol{z}_t^\top \boldsymbol{\theta}^*]} = \frac{\boldsymbol{z}_t^\top \boldsymbol{\theta}^*}{\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}}}$. At $h = \|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}} \sqrt{2\log\left(\frac{4(k-1)}{3k}\right)} \sin\left(\frac{1}{k}\left(\beta_t^\pi \wedge \frac{\pi}{2}\right)\right)$,

$$\begin{aligned}
E\left[R_t^\pi(\boldsymbol{\theta}^*, F)|\mathcal{H}_t\right] &\geq \frac{\beta_t^\pi}{4\pi}\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}} \sqrt{2\log\left(\frac{4(k-1)}{3k}\right)} \sin\left(\frac{1}{k}\left(\beta_t^\pi \wedge \frac{\pi}{2}\right)\right) \\
&= \Omega\left(\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}} \sin^2\left(\frac{\beta_t^\pi}{2}\right)\right) \\
&= \Omega\left(\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}} \cdot \left\|\frac{\boldsymbol{\theta}^*}{\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}}} - \frac{\boldsymbol{\theta}_t^\pi}{\|\boldsymbol{\theta}_t^\pi\|_{\boldsymbol{\Sigma}}}\right\|_{\boldsymbol{\Sigma}}^2\right),
\end{aligned}$$

for $k > 4$, where the first equality follows as $\frac{\beta_t^\pi \sin\left(\frac{\beta_t^\pi}{k}\right)}{\sin^2\left(\frac{\beta_t^\pi}{2}\right)} > 0$ is bounded from below by a positive constant as $\beta_t^\pi \in [0, \pi]$. Taking expectation with respect to $\mathcal{H}_t$ yields the result. □

Proof of Proposition 3.3. Suppose that $\bar{\boldsymbol{z}}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ and $\boldsymbol{z}_t = \bar{\boldsymbol{z}}_t | \|\bar{\boldsymbol{z}}_t\|_2 \leq M$. Since $E\left(\|\bar{\boldsymbol{z}}_t\|_2\right) \leq \frac{M}{2}$, we have $P\left(\|\bar{\boldsymbol{z}}_t\|_2 > M\right) \leq \frac{1}{2}$ by Markov's inequality. By Lemma 3.1 and the symmetry of the normal distribution, we have

$$\begin{aligned}
E\left[R_t^\pi(\boldsymbol{\theta}^*, F)|\mathcal{H}_t\right] &\geq \frac{h}{2} P\left(\left(\bar{\boldsymbol{z}}_t^\top \boldsymbol{\theta}^*\right)\left(\bar{\boldsymbol{z}}_t^\top \boldsymbol{\theta}_t^\pi\right) < 0 \middle| \|\bar{\boldsymbol{z}}_t\|_2 \leq M, \mathcal{H}_t\right) \\
&\quad - hP\left(\left|\bar{\boldsymbol{z}}_t^\top \boldsymbol{\theta}^*\right| \leq h,\ \mathrm{sgn}(\bar{\boldsymbol{z}}_t^\top \boldsymbol{\theta}^*) \neq \mathrm{sgn}(\bar{\boldsymbol{z}}_t^\top \boldsymbol{\theta}_t^\pi) \middle| \|\bar{\boldsymbol{z}}_t\|_2 \leq M, \mathcal{H}_t\right) \\
&= hP\left(\bar{\boldsymbol{z}}_t^\top \boldsymbol{\theta}_t^\pi < 0 < \bar{\boldsymbol{z}}_t^\top \boldsymbol{\theta}^* \middle| \|\bar{\boldsymbol{z}}_t\|_2 \leq M, \mathcal{H}_t\right) - 2hP\left(\bar{\boldsymbol{z}}_t^\top \boldsymbol{\theta}_t^\pi < 0 < \bar{\boldsymbol{z}}_t^\top \boldsymbol{\theta}^* < h \middle| \|\bar{\boldsymbol{z}}_t\|_2 \leq M, \mathcal{H}_t\right) \\
&\geq h\left[\frac{\beta_t^\pi}{2\pi} - \frac{2P\left(\bar{\boldsymbol{z}}_t^\top \boldsymbol{\theta}_t^\pi < 0 < \bar{\boldsymbol{z}}_t^\top \boldsymbol{\theta}^* < h \middle| \mathcal{H}_t\right)}{P\left(\|\bar{\boldsymbol{z}}_t\|_2 \leq M\right)}\right] \\
&\geq h\left[\frac{\beta_t^\pi}{2\pi} - \frac{2\beta_t^\pi}{\pi}\left(1 - \frac{k-1}{k} e^{-\frac{h^2}{2\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}}^2 \sin^2\left(\frac{1}{k}\left(\beta \wedge \frac{\pi}{2}\right)\right)}}\right)\right] \\
&= \frac{2h\beta_t^\pi}{\pi}\left[\frac{k-1}{k} e^{-\frac{h^2}{2\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}}^2 \sin^2\left(\frac{1}{k}\left(\beta \wedge \frac{\pi}{2}\right)\right)}} - \frac{3}{4}\right],
\end{aligned}$$

for $k > 1$ and $h > 0$, where $\beta_t^\pi = \arccos\left(\frac{(\boldsymbol{\theta}_t^\pi)^\top \boldsymbol{\Sigma}\boldsymbol{\theta}^*}{\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}}}\right)$ and the last inequality follows from Lemma A.1. At $h = \|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}} \sqrt{2\log\left(\frac{8(k-1)}{7k}\right)} \sin\left(\frac{1}{k}\left(\beta_t^\pi \wedge \frac{\pi}{2}\right)\right)$,

$$\begin{aligned}
E\left[R_t^\pi(\boldsymbol{\theta}^*, F)|\mathcal{H}_t\right] &\geq \frac{\beta_t^\pi}{4\pi}\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}} \sqrt{2\log\left(\frac{8(k-1)}{7k}\right)} \sin\left(\frac{1}{k}\left(\beta_t^\pi \wedge \frac{\pi}{2}\right)\right) \\
&= \Omega\left(\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}} \sin^2\left(\frac{\beta_t^\pi}{2}\right)\right) = \Omega\left(\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}} \cdot \left\|\frac{\boldsymbol{\theta}^*}{\|\boldsymbol{\theta}^*\|_{\boldsymbol{\Sigma}}} - \frac{\boldsymbol{\theta}_t^\pi}{\|\boldsymbol{\theta}_t^\pi\|_{\boldsymbol{\Sigma}}}\right\|_{\boldsymbol{\Sigma}}^2\right),
\end{aligned}$$

for $k > 8$, where the first equality follows as $\frac{\beta_t^\pi \sin\left(\frac{\beta_t^\pi}{k}\right)}{\sin^2\left(\frac{\beta_t^\pi}{2}\right)} > 0$ is bounded from below by a positive constant as $\beta_t^\pi \in [0, \pi]$. Taking expectation with respect to $\mathcal{H}_t$ yields the result. □

The rest of the proofs are essentially evaluating the lower bound in the previous theorems, with the aid of existing literature in estimation theory or simply following the routine therein.

Proof of Theorem 3.4. Given that $\boldsymbol{z}_t^i$ is non-zero, which happens with probability $\frac{1}{\lfloor d/2 \rfloor}$, the problem in period t is essentially equivalent to a two-dimensional one with $F = \mathcal{N}\left(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_2\right)$ and true parameter $\boldsymbol{\theta}^{*i}$. Thus, by Theorem 3.2, the conditional expected regret in period t

$$E\left[R_t^\pi(\boldsymbol{\theta}^*, F)\middle|\mathcal{H}_t, \boldsymbol{z}_t^i \neq 0\right] = \Omega\left(\sigma\left\|\boldsymbol{\theta}^{*i}\right\|_2 \cdot E\left[\left\|\frac{\boldsymbol{\theta}^{*i}}{\|\boldsymbol{\theta}^{*i}\|_2} - \frac{(\boldsymbol{\theta}_t^\pi)^i}{\|(\boldsymbol{\theta}_t^\pi)^i\|_2}\right\|_2^2\right]\right).$$

Taking expectation with respect to the conditions yield the results. □

Proof of Proposition 3.5. Given that $\boldsymbol{z}_t^i \neq \mathbf{0}$ occurs with probability $\frac{1}{\lfloor d/2 \rfloor}$, the problem in period t is essentially a two-dimensional one with a truncated $\mathcal{N}\left(\mathbf{0}, \sigma^2 \boldsymbol{I}_2\right)$ random vector $\boldsymbol{z}_t^i$ and $\boldsymbol{\theta}^{*i}$. Then, we can apply Proposition 3.3 to obtain that

$$E\left[R_t^\pi(\boldsymbol{\theta}^*, F)\middle|\mathcal{H}_t, \boldsymbol{z}_t^i \neq 0\right] \quad = \quad \Omega\left(\sigma\left\|\boldsymbol{\theta}^{*i}\right\|_2 \cdot E\left[\left\|\frac{\boldsymbol{\theta}^{*i}}{\|\boldsymbol{\theta}^{*i}\|_2} - \frac{(\boldsymbol{\theta}_t^\pi)^i}{\|(\boldsymbol{\theta}_t^\pi)^i\|_2}\right\|_2^2\right]\right).$$

□

Proof of Proposition 3.7. By Theorem 3.6, it suffices to find a lower bound of $\inf_{\boldsymbol{\theta}_t(n) \in \hat{\boldsymbol{\Theta}}_t(n)}\left\{E_\nu\left[\|\boldsymbol{\theta}^*(n)\|_{\boldsymbol{\Sigma}} \cdot \left\|\frac{\boldsymbol{\theta}^*(n)}{\|\boldsymbol{\theta}^*(n)\|_{\boldsymbol{\Sigma}}} - \frac{\boldsymbol{\theta}_t(n)}{\|\boldsymbol{\theta}_t(n)\|_{\boldsymbol{\Sigma}}}\right\|_{\boldsymbol{\Sigma}}^2\right]\right\}$. For that purpose, we may assume that all the elements in $\boldsymbol{\theta}^*$ except those in $\boldsymbol{\theta}^{*i}$ are known as it only reduces the estimation error. In addition to $\mathcal{H}_t$, the estimators in $\hat{\boldsymbol{\Theta}}_t(n)$ are also allowed access to the reward in period s from the unique unpulled arm a_s' whose binary representation is different from that of the selected arm only in the nth digit, i.e., $b_n(a_s') \neq b_n(a_s^\pi)$ while $b_{n'}(a_s') = b_{n'}(a_s^\pi)$ for all $n' \neq n$, for $1 \leq s < t$. Then, we can obtain $\boldsymbol{x}_{sa}^\top(n)\boldsymbol{\theta}^*(n) + \xi_{sa} = y_{sa} - \sum_{n' \neq n} \boldsymbol{x}_{ta}^\top(n')\boldsymbol{\theta}^*(n')$ for $a = a_s^\pi, a_s'$, $1 \leq s < t$, and $\{\boldsymbol{x}_{s,a_s^\pi}(n), \boldsymbol{x}_{s,a_s'}(n)\} = \left\{\boldsymbol{x}_{s,1}(n), \boldsymbol{x}_{s,2^{\lfloor \log_2(K) \rfloor}}(n)\right\}$ is independent of all the contexts and rewards in periods prior to s by (2)–(3). Thus, we can apply a similar argument as for the special case in Section 3.1.1 to obtain that, for any $r > 0$, there exists a distribution ν such that $P_\nu\left(\|\boldsymbol{\theta}^*(n)\|_2 = r\right) = 1$ and $\inf_{\boldsymbol{\theta}_t(n) \in \hat{\boldsymbol{\Theta}}_t(n)}\left\{E_\nu\left[\|\boldsymbol{\theta}^*(n)\|_{\boldsymbol{\Sigma}} \cdot \left\|\frac{\boldsymbol{\theta}^*(n)}{\|\boldsymbol{\theta}^*(n)\|_{\boldsymbol{\Sigma}}} - \frac{\boldsymbol{\theta}_t(i)}{\|\boldsymbol{\theta}_t(i)\|_{\boldsymbol{\Sigma}}}\right\|_{\boldsymbol{\Sigma}}^2\right]\right\} = \Omega\left(\frac{r}{\sqrt{d}}\left(\frac{1}{2} - \frac{32r^2 t \log_2(K)}{d^2}\right)\right)$ when $d \gg 1$. Therefore,

$$\inf_{\pi \in \Pi}\left\{\sup_{\boldsymbol{\theta}^* \in \boldsymbol{\Theta}^*}\left\{R^\pi\left(\boldsymbol{\theta}^*, F_K\right)\right\}\right\} = \Omega\left(\frac{r\log_2(K)}{\sqrt{d}}\sum_{t=1}^T\left(\frac{1}{2} - \frac{32r^2 t\log_2(K)}{d^2}\right)\right),$$

which is $\Omega\left(\max\left\{\sqrt{dT\log_2(K)}, d^{3/2}\right\}\right)$ at $r = \min\left\{\frac{d}{4\sqrt{T\log_2(K)}}, \frac{1}{\sqrt{\log_2(K)}}\right\}$. □

Proof of Proposition 4.1. Let $\hat{\boldsymbol{\theta}}_t$ be the minimizer of $\inf_{\boldsymbol{\theta}_t \in \hat{\boldsymbol{\Theta}}_t}\left\{\sum_{i=1}^{\lfloor d/2 \rfloor} E_\nu\left[\left\|\boldsymbol{\theta}^{*i} - r\boldsymbol{\theta}_t^i\right\|_2^2\right]\right\}$ and $N_t^i = \sum_{s=1}^{t-1} 1_{\{\boldsymbol{z}_s^i \neq \mathbf{0}\}}$ be the number of periods up to $t-1$ that $\boldsymbol{\theta}^{*i}$, $1 \leq i \leq \left\lfloor \frac{d}{2} \right\rfloor$, can be learnt. For any $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \mathbb{R}^d$, by Theorem 1 in (Devroye et al., 2018), the total variation distance between the distribution of $\overline{\mathcal{H}}_t$ conditioned on $\boldsymbol{\theta}^{*i} = \boldsymbol{\theta}^i$ and that conditioned on $\boldsymbol{\theta}^{*i} = \boldsymbol{\theta}'^i$ is bounded from above by $\left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2 \sqrt{N_t^i}$. Thus,

$$\begin{aligned}
&P\left(\left\|\boldsymbol{\theta}^{*i} - r\hat{\boldsymbol{\theta}}_t^i\right\|_2^2 \geq \frac{1}{4}\left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2^2 \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}, N_t^i\right) \\
&\geq P\left(\left\|\boldsymbol{\theta}^i - r\hat{\boldsymbol{\theta}}_t^i\right\|_2 \geq \left\|\boldsymbol{\theta}'^i - r\hat{\boldsymbol{\theta}}_t^i\right\|_2 \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}, N_t^i\right) \\
&\geq \frac{1}{2}\left[P\left(\left\|\boldsymbol{\theta}^i - r\hat{\boldsymbol{\theta}}_t^i\right\|_2 \geq \left\|\boldsymbol{\theta}'^i - r\hat{\boldsymbol{\theta}}_t^i\right\|_2 \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}, N_t^i\right) + 1 - P\left(\left\|\boldsymbol{\theta}^i - r\hat{\boldsymbol{\theta}}_t^i\right\|_2 \leq \left\|\boldsymbol{\theta}'^i - r\hat{\boldsymbol{\theta}}_t^i\right\|_2 \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}, N_t^i\right)\right] \\
&= \frac{1}{2}\left[1 + P\left(\left\|\boldsymbol{\theta}^i - r\hat{\boldsymbol{\theta}}_t^i\right\|_2 \geq \left\|\boldsymbol{\theta}'^i - r\hat{\boldsymbol{\theta}}_t^i\right\|_2 \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}, N_t^i\right) - P\left(\left\|\boldsymbol{\theta}^i - r\hat{\boldsymbol{\theta}}_t^i\right\|_2 \geq \left\|\boldsymbol{\theta}'^i - r\hat{\boldsymbol{\theta}}_t^i\right\|_2 \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}', N_t^i\right)\right] \\
&\geq \frac{1}{2}\left(1 - \left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2 \sqrt{N_t^i}\right),
\end{aligned}$$

where the equality follows from the symmetry of the distribution of $\overline{\mathcal{H}}_t$ and ν. Since, for given $\boldsymbol{\theta} \in \boldsymbol{\Theta}_t^*$ and $q \in [0, r^2]$, there exists $\boldsymbol{\theta}' \in \boldsymbol{\Theta}^*$ such that $\left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2 = 2\sqrt{q}$, we can evaluate $E\left[\left\|\boldsymbol{\theta}^{*i} - r\hat{\boldsymbol{\theta}}_t^i\right\|_2^2\right]$ as

$$\begin{aligned}
& E\left[E\left(\left\|\boldsymbol{\theta}^{*i} - r\boldsymbol{\theta}_t^i\right\|_2^2 \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}, N_t^i\right)\right] \\
\geq\ & E\left[\int_0^{r^2} P\left(\left\|\boldsymbol{\theta}_t^i - \boldsymbol{\theta}^{*i}\right\|_2^2 \geq q \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}, N_t^i\right) dq\right] \\
\geq\ & \frac{1}{2}E\left[\int_0^{r^2} \left(1 - 2\sqrt{qN_t^i}\right)^+ dq\right] \\
\geq\ & \frac{1}{2}\int_0^{r^2} E\left[\left(\frac{1}{2} - 2qN_t^i\right)^+\right] dq \\
\geq\ & \frac{1}{2}\int_0^{r^2} \left(\frac{1}{2} - 2qE\left[N_t^i\right]\right)^+ dq = \Omega\left(\min\left\{r^2, \frac{d}{t}\right\}\right).
\end{aligned}$$

By Proposition 3.4, $\inf_{\pi \in \Pi}\left\{\sup_{\boldsymbol{\theta}^* \in \boldsymbol{\Theta}^*}\{R^\pi(\boldsymbol{\theta}^*, F)\}\right\} = \frac{\sigma}{dr}\sum_{t=1}^{T}\sum_{i=1}^{\lfloor d/2 \rfloor}\Omega\left(\min\left\{r^2, \frac{d}{t}\right\}\right) = \sum_{t=1}^{T}\Omega\left(\sigma \min\left\{r, \frac{d}{rt}\right\}\right)$. □

Proof of Proposition 4.3. We now calculate the bound in Theorem 3.5. Letting $\hat{\boldsymbol{\theta}}_t$ be the minimizer of $\inf_{\boldsymbol{\theta}_t \in \hat{\boldsymbol{\Theta}}_t} E_\nu\left[\|\boldsymbol{\theta}^* - r\boldsymbol{\theta}_t\|_2^2\right]$ and $N_t^i = \sum_{s=1}^{t-1} 1_{\{\boldsymbol{z}_s^i \neq \boldsymbol{0}\}}$, by Lemma 6.1 in (Karwa & Vadhan, 2017), we have that, for any $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \mathbb{R}^d$,

$$\begin{aligned}
& P\left(\left\|\hat{\boldsymbol{\theta}}_t^i - \boldsymbol{\theta}^i\right\|_2 \geq \left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2 \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}, N_t^i, \mathcal{H}_t\right) \\
& \geq e^{-6N_i^t\varepsilon\left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2} P\left(\left\|\hat{\boldsymbol{\theta}}_t^i - \boldsymbol{\theta}^i\right\|_2 \geq \left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2 \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}', N_t^i, \mathcal{H}_t\right) - \delta\left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2 N_i^t \\
& \geq e^{-6N_i^t\varepsilon\left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2}\left[1 - P\left(\left\|\hat{\boldsymbol{\theta}}_t^i - \boldsymbol{\theta}^i\right\|_2 \geq \left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2 \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}, N_t^i, \mathcal{H}_t\right)\right] - \delta\left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2 N_i^t,
\end{aligned}$$

where the last inequality follows from symmetry, implying that

$$P\left(\left\|\hat{\boldsymbol{\theta}}_t^i - \boldsymbol{\theta}^i\right\|_2 \geq \left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2 \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}, N_t^i, \mathcal{H}_t\right) \geq \frac{1}{2}\left(e^{-6N_i^t\varepsilon\left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2} - \delta\left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2 N_i^t\right).$$

Since, for given $\boldsymbol{\theta} \in \boldsymbol{\Theta}_r^*$ and $q \in [0, r^2]$, there exists $\boldsymbol{\theta}' \in \boldsymbol{\Theta}^*$ such that $\left\|\boldsymbol{\theta}^i - \boldsymbol{\theta}'^i\right\|_2 = 2\sqrt{q}$, we have

$$\begin{aligned}
E\left[\|\boldsymbol{\theta}_t - \boldsymbol{\theta}^*\|_2^2\right] &= \sum_{i=1}^{\lfloor d/2 \rfloor} E\left[E\left(\left\|\boldsymbol{\theta}_t^i - \boldsymbol{\theta}^{*i}\right\|_2^2 \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}, N_t^i, \mathcal{H}_t\right)\right] \\
&\geq \sum_{i=1}^{d/2} E\left[\int_0^{r^2} P\left(\left\|\boldsymbol{\theta}_t^i - \boldsymbol{\theta}^{*i}\right\|_2^2 \geq q \middle| \boldsymbol{\theta}^* = \boldsymbol{\theta}, N_t^i\right) dq\right] \\
&\geq \frac{1}{2}\sum_{i=1}^{d/2} E\left[\int_0^{r^2}\left(e^{-12N_i^t\varepsilon\sqrt{q}} - 2\delta N_i^t\sqrt{q}\right)^+ dq\right] \\
&\geq \frac{1}{2}\sum_{i=1}^{d/2}\int_0^{r^2}\left(e^{-12E\left[N_i^t\right]\varepsilon\sqrt{q}} - 2\delta E\left[N_i^t\right]\sqrt{q}\right)^+ dq \\
&= \Omega\left(d\min\left\{r^2, \frac{d^2}{t^2(\varepsilon + \delta)^2}\right\}\right).
\end{aligned}$$

The Proposition follow from Theorem 3.4 □

Proof of Proposition 4.4. Under the given choice of $(F, \nu, \hat{\boldsymbol{\Theta}}_t)$, the lower bound in Theorem 3.2,

$$\begin{aligned}\inf_{\pi\in\Pi}\left\{\sup_{\boldsymbol{\theta}^*\in\boldsymbol{\Theta}^*}\left\{R^\pi\left(\boldsymbol{\theta}^*, F\right)\right\}\right\} &= \Omega\left(\sum_{t=1}^T \inf_{\boldsymbol{\theta}_t\in\hat{\boldsymbol{\Theta}}_t}\left\{E_\nu\left[\frac{\left\|\boldsymbol{\theta}^* - \|\boldsymbol{\theta}^*\|_2\cdot\boldsymbol{\theta}_t\right\|_2^2}{\|\boldsymbol{\theta}^*\|_2}\right]\right\}\right)\\ &= \Omega\left(\frac{1}{r}\sum_{t=1}^T \inf_{\boldsymbol{\theta}_t\in\hat{\boldsymbol{\Theta}}_t}\left\{E_\nu\left[\|\boldsymbol{\theta}^* - r\boldsymbol{\theta}_t\|_2^2\right]\right\}\right)\\ &= \Omega\left(\frac{1}{r}\sum_{t=1}^T \inf_{\boldsymbol{\theta}_t\in\hat{\boldsymbol{\Theta}}_t}\left\{E_\nu\left[\|\boldsymbol{\theta}^* - \boldsymbol{\theta}_t\|_2^2\right]\right\}\right).\end{aligned}$$

It remains to find a lower bound for $E_\nu\left[\|\boldsymbol{\theta}^* - \boldsymbol{\theta}_t\|_2^2\right]$. (Wang & Xu, 2019) showed that the squared ℓ_2-risk for sparse linear regressions under an ε-LDP constraint has a lower bound of order $\Omega\left(\frac{d\log d}{t\varepsilon^2}\right)$ when $(\boldsymbol{x}_{ta}, y_{ta})$ can only take discrete values on the hamming cube $\{1,-1\}^{d+1}$. Since our $(\boldsymbol{x}_{ta}, y_{ta})$ take continuous values, we will provide a continuous version of their proof to reach the same lower bound as follows.

Denote by $L^\infty\left(\mathbb{R}^n\right)$ the space of essentially bounded measurable functions on $\mathbb{R}^n$. By Proposition 2 in (Duchi et al., 2018),

$$\inf_{\boldsymbol{\theta}_t\in\hat{\boldsymbol{\Theta}}_t}\left\{E_\nu\left[\|\boldsymbol{\theta}^* - \boldsymbol{\theta}_t\|_2^2\right]\right\} = \Omega\left(r^2\left\{1 - \frac{t\varepsilon^2}{\log(d)}\left[\frac{1}{2d}\sup_{\gamma\in\mathbb{B}_\infty}\left\{\sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_r^*}\Psi_{\boldsymbol{\theta}}^2(\gamma)\right\}\right] - \frac{\log(2)}{\log(d)}\right\}\right),$$

where $\mathbb{B}_\infty = \left\{\gamma\in L^\infty\left(\mathbb{R}^{2d+2}\right) : \|\gamma\|_\infty \le 1\right\}$ is the centered unit ball in $L^\infty(\mathbb{R}^{2d+2})$ and

$$\Psi_{\boldsymbol{\theta}}(\gamma) = E\left[\gamma\left(\boldsymbol{x}_{t1}, \boldsymbol{x}_{t2}, y_{t1}, y_{t2}\right)|\boldsymbol{\theta}^* = \boldsymbol{\theta}\right] - E_\nu\left[E\left[\gamma\left(\boldsymbol{x}_{t1}, \boldsymbol{x}_{t2}, y_{t1}, y_{t2}\right)|\boldsymbol{\theta}^*\right]\right].$$

In Lemma A.2 below, we provide an upper bound for $\sup_{\gamma\in\mathbb{B}_\infty}\left\{\sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_r^*}\Psi_{\boldsymbol{\theta}}^2(\gamma)\right\}$. Thus, $\inf_{\pi\in\Pi}\left\{\sup_{\boldsymbol{\theta}^*\in\boldsymbol{\Theta}^*}\left\{R^\pi\left(\boldsymbol{\theta}^*, F\right)\right\}\right\} = \Omega\left(r\left[1 - \frac{Tr^2\varepsilon^2}{d(1-2r^2)\log(d)} - \frac{\log(2)}{\log(d)}\right]\right)$. For a general K, following a similar argument as above, we have $\inf_{\pi\in\Pi}\left\{\sup_{\boldsymbol{\theta}^*\in\boldsymbol{\Theta}^*}\left\{R^\pi\left(\boldsymbol{\theta}^*, F\right)\right\}\right\} = \Omega\left(r\log_2(K)\left[1 - \frac{T\log_2(K)r^2\varepsilon^2}{d(1-2r^2)\log(d/\log_2(K))} - \frac{\log(2)}{\log(d/\log_2(K))}\right]\right)$. □

Lemma A.2. $\sup_{\gamma\in\mathbb{B}_\infty}\left\{\sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_r^*}\Psi_{\boldsymbol{\theta}}^2(\gamma)\right\} \le \frac{2r^2}{1-2r^2}$.

Proof of Lemma A.2. Let $\phi(y|\mu) = \frac{1}{\sqrt{2\pi}}e^{-\frac{(y-\mu)^2}{2}}$ denote the density of a $\mathcal{N}(\mu, 1)$ random variable. For $(\boldsymbol{x}, \boldsymbol{x}', y, y') \in \mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}$, define

$$G_{\boldsymbol{\theta}}(\boldsymbol{x}, \boldsymbol{x}', y, y') := \phi\left(y|\boldsymbol{x}^\top\boldsymbol{\theta}\right)\phi\left(y'|\boldsymbol{x}'^\top\boldsymbol{\theta}\right) - \frac{1}{2d}\sum_{\boldsymbol{\theta}'\in\boldsymbol{\Theta}_r^*}\phi\left(y|\boldsymbol{x}^\top\boldsymbol{\theta}'\right)\phi\left(y'|\boldsymbol{x}'^\top\boldsymbol{\theta}'\right),$$

and let $\boldsymbol{G}(\boldsymbol{x}, \boldsymbol{x}', y, y')$ be the $2d$-dimensional column vector defined by

$$\boldsymbol{G}(\boldsymbol{x}, \boldsymbol{x}', y, y') := \left(G_{r\boldsymbol{e}_1}(\boldsymbol{x}, \boldsymbol{x}', y, y'), G_{-r\boldsymbol{e}_1}(\boldsymbol{x}, \boldsymbol{x}', y, y'), \ldots, G_{r\boldsymbol{e}_d}(\boldsymbol{x}, \boldsymbol{x}', y, y'), G_{-r\boldsymbol{e}_d}(\boldsymbol{x}, \boldsymbol{x}', y, y')\right)^\top.$$

By the Cauchy-Schwartz inequality,

$$\begin{aligned}
&\sup_{\gamma\in\mathbb{B}_\infty}\left\{\sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_r^*}\Psi_{\boldsymbol{\theta}}^2(\gamma)\right\}\\
&=\sup_{\gamma\in\mathbb{B}_\infty}\left\{\sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_r^*}\left(\int_{y,y'}E\left[\gamma(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')G_{\boldsymbol{\theta}}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\right]dydy'\right)^2\right\}\\
&=\sup_{\gamma\in\mathbb{B}_\infty}\left\{\sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_r^*}\left(\int_{y,y'}\frac{E\left[\gamma(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')G_{\boldsymbol{\theta}}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\right]}{\sqrt{\phi(y|0)\phi(y'|0)}}\sqrt{\phi(y|0)\phi(y'|0)}dydy'\right)^2\right\}\\
&\le\sup_{\gamma\in\mathbb{B}_\infty}\left\{\sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_r^*}\int_{y,y'}\left(\frac{E\left[\gamma(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')G_{\boldsymbol{\theta}}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\right]}{\sqrt{\phi(y|0)\phi(y'|0)}}\right)^2dydy'\int_{y,y'}\phi(y|0)\phi(y'|0)dydy'\right\}\\
&=\sup_{\gamma\in\mathbb{B}_\infty}\left\{\sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_r^*}\int_{y,y'}E^2\left[\gamma(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')G_{\boldsymbol{\theta}}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\right]\frac{dydy'}{\phi(y|0)\phi(y'|0)}\right\}\\
&\le\int_{y,y'}\sup_{\gamma\in\mathbb{B}_\infty}\left\{\sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_r^*}E^2\left[\gamma(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')G_{\boldsymbol{\theta}}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\right]\right\}\frac{dydy'}{\phi(y|0)\phi(y'|0)}\\
&=\int_{y,y'}\sup_{\gamma\in\mathbb{B}_\infty}\left\{\left\|E\left[\gamma(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\boldsymbol{G}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\right]\right\|_2^2\right\}\frac{dydy'}{\phi(y|0)\phi(y'|0)}\\
&=\int_{y,y'}\sup_{\substack{\gamma\in\mathbb{B}_\infty,\\ \boldsymbol{u}\in\mathbb{R}^{2d},\|\boldsymbol{u}\|_2\le1}}\left\{E^2\left[\gamma(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\boldsymbol{u}^\top\boldsymbol{G}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\right]\right\}\frac{dydy'}{\phi(y|0)\phi(y'|0)}\\
&\le\int_{y,y'}\sup_{\substack{\gamma\in\mathbb{B}_\infty,\\ \boldsymbol{u}\in\mathbb{R}^{2d},\|\boldsymbol{u}\|_2\le1}}\left\{E\left[\gamma^2(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\right]E\left[\left(\boldsymbol{u}^\top\boldsymbol{G}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\right)^2\right]\right\}\frac{dydy'}{\phi(y|0)\phi(y'|0)}\\
&\le\int_{y,y'}\sup_{\boldsymbol{u}\in\mathbb{R}^{2d},\|\boldsymbol{u}\|_2\le1}\left\{\boldsymbol{u}^\top E\left[\boldsymbol{G}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\boldsymbol{G}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')^\top\right]\boldsymbol{u}\right\}\frac{dydy'}{\phi(y|0)\phi(y'|0)}.
\end{aligned}$$

Note that $\sup_{\boldsymbol{u}\in\mathbb{R}^{2d},\|\boldsymbol{u}\|_2\le1}\left\{\boldsymbol{u}^\top E\left[\boldsymbol{G}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\boldsymbol{G}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')^\top\right]\boldsymbol{u}\right\}$ is equal to the largest eigenvalue of $E\left[\boldsymbol{G}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\boldsymbol{G}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')^\top\right]$. For given $\boldsymbol{\theta},\boldsymbol{\theta}'\in\boldsymbol{\Theta}_r^*=\{r\boldsymbol{e}_1,-r\boldsymbol{e}_1,\dots,r\boldsymbol{e}_d,-r\boldsymbol{e}_d\}$,

$$\begin{aligned}
&E\left[G_{\boldsymbol{\theta}}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')G_{\boldsymbol{\theta}'}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\right]\\
&=E\left[\phi(y|\boldsymbol{x}_{t1}^\top\boldsymbol{\theta})\phi(y|\boldsymbol{x}_{t1}^\top\boldsymbol{\theta}')\phi(y'|\boldsymbol{x}_{t2}^\top\boldsymbol{\theta})\phi(y'|\boldsymbol{x}_{t2}^\top\boldsymbol{\theta}')\right]+\frac{1}{4d^2}\sum_{\boldsymbol{\beta},\boldsymbol{\beta}'\in\boldsymbol{\Theta}_r^*}E\left[\phi(y|\boldsymbol{x}_{t1}^\top\boldsymbol{\beta})\phi(y|\boldsymbol{x}_{t1}^\top\boldsymbol{\beta}')\phi(y'|\boldsymbol{x}_{t2}^\top\boldsymbol{\beta})\phi(y'|\boldsymbol{x}_{t2}^\top\boldsymbol{\beta}')\right]\\
&\quad-\frac{1}{2d}\sum_{\boldsymbol{\beta}\in\boldsymbol{\Theta}_r^*}E\left[\phi(y|\boldsymbol{x}_{t1}^\top\boldsymbol{\beta})\phi(y'|\boldsymbol{x}_{t2}^\top\boldsymbol{\beta})\left(\phi(y|\boldsymbol{x}_{t1}^\top\boldsymbol{\theta})\phi(y'|\boldsymbol{x}_{t2}^\top\boldsymbol{\theta})+\phi(y|\boldsymbol{x}_{t1}^\top\boldsymbol{\theta}')\phi(y'|\boldsymbol{x}_{t2}^\top\boldsymbol{\theta}')\right)\right]\\
&=E\left[\phi(y|\boldsymbol{x}_{t1}^\top\boldsymbol{\theta})\phi(y|\boldsymbol{x}_{t1}^\top\boldsymbol{\theta}')\phi(y'|\boldsymbol{x}_{t2}^\top\boldsymbol{\theta})\phi(y'|\boldsymbol{x}_{t2}^\top\boldsymbol{\theta}')\right]-\frac{1}{2d}\frac{e^{-\frac{y^2+y'^2}{1+2r^2}}}{1+2r^2}-\left(1-\frac{1}{d}\right)\frac{e^{-\frac{y^2+y'^2}{1+r^2}}}{1+r^2}-\frac{1}{2d}\frac{e^{-(y^2+y'^2)}}{1+2r^2}\\
&=\begin{cases}
\left(1-\frac{1}{2d}\right)\frac{e^{-\frac{y^2+y'^2}{1+2r^2}}}{1+2r^2}-\left(1-\frac{1}{d}\right)\frac{e^{-\frac{y^2+y'^2}{1+r^2}}}{1+r^2}-\frac{1}{2d}\frac{e^{-(y^2+y'^2)}}{1+2r^2}\triangleq H_+, & \text{if }\boldsymbol{\theta}=\boldsymbol{\theta}',\\
-\frac{1}{2d}\frac{e^{-\frac{y^2+y'^2}{1+2r^2}}}{1+2r^2}-\left(1-\frac{1}{d}\right)\frac{e^{-\frac{y^2+y'^2}{1+r^2}}}{1+r^2}+\left(1-\frac{1}{2d}\right)\frac{e^{-(y^2+y'^2)}}{1+2r^2}\triangleq H_-, & \text{if }\boldsymbol{\theta}=-\boldsymbol{\theta}',\\
-\frac{1}{2d}\frac{e^{-\frac{y^2+y'^2}{1+2r^2}}}{1+2r^2}-\left(1-\frac{1}{d}\right)\frac{e^{-\frac{y^2+y'^2}{1+r^2}}}{1+r^2}+\left(1-\frac{1}{2d}\right)\frac{e^{-(y^2+y'^2)}}{1+2r^2}=-\frac{H_++H_-}{2d-2}, & \text{if }\boldsymbol{\theta}\ne\pm\boldsymbol{\theta}'.
\end{cases}
\end{aligned}$$

Therefore,

$$E\left[\boldsymbol{G}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\boldsymbol{G}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')^\top\right] = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} & \cdots & \boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{A} & \cdots & \boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{B} & \boldsymbol{B} & \cdots & \boldsymbol{A} \end{pmatrix},$$

where

$$\boldsymbol{A} = \begin{pmatrix} H_+ & H_- \\ H_- & H_+ \end{pmatrix} \text{ and } \boldsymbol{B} = -\frac{H_+ + H_-}{2d-2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

It can then be easily verified that $E\left[\boldsymbol{G}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')\boldsymbol{G}(\boldsymbol{x}_{t1},\boldsymbol{x}_{t2},y,y')^\top\right]$ has the following three eigenvalues:

1. $H_+ - H_-$, with a d-dimensional eigenspace spanned by $(1,-1,0,\dots,0), (0,0,1,-1,0,\dots,0), \cdots,$ $(0,0,\dots,0,1,-1)$.
2. $\frac{2d(H_+ + H_-)}{2d-2}$, with a d-dimensional eigenspace spanned by $(1,1,-1,-1,0,\dots,0)$, $(0,0,1,1,-1,-1,0,\dots,0), \cdots, (0,\dots,0,1,1,-1,-1)$.
3. 0, with a 1-dimensional eigenspace spanned by $(1,1,\dots,1)$.

Since the eigenspaces of the three eigenvalues listed above span the whole space $\mathbb{R}^{2d}$, they are the only eigenvalues with the largest being $H_+ - H_- = \frac{e^{-\frac{y^2+y'^2}{1+2r^2}} - e^{-(y^2+y'^2)}}{1+2r^2}$. Thus,

$$\sup_{\gamma\in\mathbb{B}_\infty}\left\{\sum_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_r^*}\Psi_{\boldsymbol{\theta}}^2(\gamma)\right\} \le \int_{y,y'} \frac{e^{-\frac{y^2+y'^2}{1+2r^2}} - e^{-(y^2+y'^2)}}{1+2r^2}\frac{dydy'}{\phi(y|0)\phi(y'|0)} \le \frac{2r^2}{1-2r^2}.$$

□