New Classes of the Greedy-Applicable Arm Feature Distributions in the Sparse Linear Bandit Problem

Koji Ichikawa[1 2], **Shinji Ito**[1 2 3], **Daisuke Hatano**[3], **Hanna Sumita**[4], **Takuro Fukunaga**[5], **Naonori Kakimura**[6], **Ken-ichi Kawarabayashi**[7 8]

[1]NEC Corporation
[2]National Institute of Advanced Industrial Science and Technology
[3]RIKEN AIP
[4]Tokyo Institute of Technology
[5]Chuo University
[6]Keio University
[7]National Institute of Informatics
[8]The University of Tokyo
{k_ichikawa0, i-shinji}@nec.com, daisuke.hatano@riken.jp, sumita@c.titech.ac.jp, fukunaga.07s@chuo-u.ac.jp, kakimura@math.keio.ac.jp, k_keniti@nii.ac.jp

Abstract

We consider the sparse contextual bandit problem where arm feature affects reward through the inner product of sparse parameters. Recent studies have developed sparsity-agnostic algorithms based on the greedy arm selection policy. However, the analysis of these algorithms requires strong assumptions on the arm feature distribution to ensure that the greedily selected samples are sufficiently diverse; One of the most common assumptions, relaxed symmetry, imposes approximate origin-symmetry on the distribution, which cannot allow distributions that has origin-asymmetric support. In this paper, we show that the greedy algorithm is applicable to a wider range of the arm feature distributions from two aspects. Firstly, we show that a mixture distribution that has a greedy-applicable component is also greedy-applicable. Second, we propose new distribution classes, related to Gaussian mixture, discrete, and radial distribution, for which the sample diversity is guaranteed. The proposed classes can describe distributions with origin-asymmetric support and, in conjunction with the first claim, provide theoretical guarantees of the greedy policy for a very wide range of the arm feature distributions.

1 Introduction

The contextual bandit problems are extensively investigated in various settings (Lattimore and Szepesvári 2020), with practical applications to recommendations (Li et al. 2010), clinical trials (Durand et al. 2018; Bastani and Bayati 2020), and many others (Bouneffouf and Rish 2019). The problems are sequential decision making problems where, in each round, a learner observes a set of arms with context, chooses one of them, and receives a corresponding reward. In this paper, we assume that arm features (context) for each arm are stochastically generated and that the reward is affected by the inner product of the selected arm features and unknown parameters. The problem is known to admit sublinear-regret algorithms that utilize the upper confidence bound for the arm selection criteria (Auer 2002; Dani, Hayes, and Kakade 2008; Li et al. 2010; Rusmevichientong and Tsitsiklis 2010; Chu et al. 2011; Abbasi-Yadkori, Pál, and Szepesvári 2011), but they usually allow the unknown parameters to be dense.

In the sparse linear bandits, introduced by Abbasi-Yadkori, Pál, and Szepesvári (2012) and Carpentier and Munos (2012), we consider the situations where the arm features are high-dimensional, but the unknown parameters are sparse. If we are given the sparsity of the unknown parameters in advance, the sparse linear bandits admit sublinear-regret algorithms (Abbasi-Yadkori, Pál, and Szepesvári 2012; Carpentier and Munos 2012; Bastani and Bayati 2020; Wang, Wei, and Yao 2018; Kim and Paik 2019), which outperform the linear bandit algorithms for the dense setting by exploiting the sparsity. On the other hand, for the estimation of the unknown parameters, the arm features chosen by the algorithms must be sufficiently diverse. In particular, recent algorithms (Bastani and Bayati 2020; Wang, Wei, and Yao 2018; Kim and Paik 2019) adopt the forced-sampling or uniform sampling step to guarantee the diversity of the chosen arm features. We need to know the sparsity in advance to ensure an optimal ratio between the forced or uniform sampling step and the other strategic arm selection step.

Oh, Iyengar, and Zeevi (2021) and Ariu, Abe, and Proutière (2022) recently proposed algorithms for the sparse linear bandits that work without knowing the sparsity in advance. Such sparsity-agnostic algorithms are based on the greedy arm selection policy. They showed that the greedy arm selection automatically guarantees the sample diversity under certain assumptions on the arm feature distribution. However, the existing analysis of sparsity-agnostic algorithms is of limited applicability in practice due to their strong assumption on the arm feature distribution. One of their typical assumptions is relaxed symmetry, which requires the arm feature distribution to be almost symmetric around the origin. Thus the current analysis is not applicable to the problem where arm features take only positive values, arising in practical applications such as a recommender system.

In this paper, we show that the greedy algorithm is in fact applicable to the problem with a wider range of the arm feature distributions from two aspects. Firstly, we show that, if an arm feature distribution is greedy-applicable, then so does their mixture. Here we call arm feature distribution greedy-applicable if the sample diversity is guaranteed under the greedy arm selection policy. It is also shown in the proof that a larger proportion of the greedy-applicable component in the mixture distribution yields a tighter regret upper bound, motivating the presentation of a variety of greedy-applicable distribution classes. Secondly, we propose new representational classes of the greedy-applicable distributions, related to Gaussian mixture, discrete, and radial distribution. The proposed classes can describe distributions with origin-asymmetric support. These two generalizations provide theoretical guarantees of the greedy policy for a very wide range of the arm feature distributions. Moreover, we demonstrate the usefulness of our analysis by applying it to the other cases: thresholded lasso bandit (Ariu, Abe, and Proutière 2022), combinatorial setting, and non-sparse setting (Bastani and Bayati 2020).

The organization of this paper is as follows: in Section 2, we describe related work; in Section 3, we introduce the regret analysis for the greedy policy on sparse linear bandits and the assumptions for the arm feature distribution; in Section 4, we present our theorems on the greedy-applicable distributions. Section 5 is devoted to applications of our analysis and, finally, we give a discussion and conclusion in Section 6.

2 Related Works

In the sparse linear bandit problem, Abbasi-Yadkori, Pál, and Szepesvári (2012) proposed an algorithm with the online-to-confidence-set conversion technique, giving a regret upper bound of $O(\sqrt{s_0 dT})$. Here, T is the horizon and d is the dimension of the arm features, and s_0 is the sparsity of the unknown parameter. Bastani and Bayati (2020) and Wang, Wei, and Yao (2018) dealt with a multi-parameter setting, i.e., a setting where each arm has an unknown sparse parameter and arm selection is performed for a vector of context, giving regret upper bounds of $O(s_0^2(\log dT)^2)$ and $O(s_0^2 \log d \log T)$, respectively. In the single-parameter setting, Kim and Paik (2019) showed an $O(s_0 \log(dT)\sqrt{T})$ upper bound using the doubly-robust lasso bandit approach.

In the above series of studies, their algorithms used prior knowledge of s_0 as input to the algorithm. Oh, Iyengar, and Zeevi (2021) proposed a sparsity-agnostic algorithm and, under the assumption for the arm feature distribution, called relaxed symmetry, gave a regret upper bound of $O(s_0\sqrt{\log(dT)T})$. The thresholded lasso bandit algorithm proposed by Ariu, Abe, and Proutière (2022) is another sparsity-agnostic algorithm. They showed an $O(\sqrt{s_0 T})$ upper bound under an additional assumption about the sparse positive definiteness of the arm feature distribution. However, the relaxed symmetry imposed for their regret analysis was severe. Our study shows that the greedy-based algorithm in fact is applicable to a wider class of the arm feature distributions.

Another line of research is to consider the greedy algorithm in the dense parameter settings. In Bastani, Bayati, and Khosravi (2021), the greedy algorithm was studied in the multi-parameter setting. They showed that by introducing the covariate diversity assumption, the greedy algorithm for two arms achieves the rate optimal. Kannan et al. (2018) investigated the greedy algorithm in the perturbed adversarial setting. The authors showed that even in adversarial samples, the addition of stochastic isotropic Gaussian perturbations makes the regret upper bound of $O(d\sqrt{dT})$. For sparse linear bandit, Sivakumar, Wu, and Banerjee (2020) gave an $O(\sqrt{s_0 dT})$ upper bound in the perturbed adversarial setting.

3 Preliminary

3.1 Problem Setup

In this article, we consider the following linear bandit problem: We are given a horizon T. For each round $t \in [T]$, we observe a set of K arm features of d dimensional vector $\mathcal{X}^t := \{X_1^t, \dots, X_K^t\}$, where $X_i^t \in \mathbb{R}^d$ for $i \in [K]$, generated by an arm feature distribution $P(\mathcal{X}^t)$. For each round t, we select one of these arms with index $a_t \in [K]$ and receive a corresponding reward $r_t \in \mathbb{R}$.

In the sparse linear contextual bandit, the observed rewards are modeled as the inner product of the selected arm feature $X_{a_t}^t$ and an unknown sparse parameter $\beta^* \in \mathbb{R}^d$ with sparsity $\|\beta^*\|_0 = s_0$:

$$r_t := X_{a_t}^{t\top} \beta^* + \epsilon_t \,, \tag{1}$$

where ϵ_t is a σ-sub-Gaussian noise satisfying $\mathbb{E}[e^{\lambda \epsilon_t} \mid \mathcal{F}_t] \leq e^{\lambda^2 \sigma^2/2}$ for any $\lambda \in \mathbb{R}$, with $\mathcal{F}_t$ being the σ-algebra $\sigma(\mathcal{X}^1, a_1, r_1, \dots, \mathcal{X}^{t-1}, a_{t-1}, r_{t-1}, \mathcal{X}^t, a_t)$. For later use, we here define another σ-algebra: $\mathcal{F}_t' := \sigma(\mathcal{X}^1, a_1, r_1, \dots, \mathcal{X}^t, a_t, r_t)$. The problem is to minimize the following expected regret:

$$R(T) = \sum_{t=1}^{T} \mathbb{E}[r_t^* - r_t] \,, \tag{2}$$

under selection criteria $\{a_1, \dots, a_T\}$. Here r_t^* is the reward for the optimal arm choice in round t.

Below, we make the following standard assumptions about the bound for the arm feature $\mathcal{X} = \{X_1, \dots, X_K\} \sim P(\mathcal{X})$ and β^*.

Assumption 1. *The arm feature X_i for each $i \in [K]$ and β^* are bounded as $\|X_i\|_\infty \leq x_{\max} < \infty$ and $\|\beta^*\|_1 \leq b < \infty$, respectively.*

3.2 LASSO Estimator

Here we introduce the theories of LASSO (Tibshirani 1996), which is employed in most papers on the sparse linear bandits. LASSO estimates the parameter β^* under the observed samples $(X_{a_1}^1, r_1, \dots, X_{a_t}^t, r_t)$ by the following L^1-regularized least square:

$$\hat{\beta}_t = \operatorname*{argmin}_{\beta} \frac{1}{t} \sum_{s=1}^{t} (r_s - X_{a_s}^{s\top} \beta)^2 + \lambda_t \|\beta\|_1 \,, \tag{3}$$

where $\lambda_t > 0$ is a hyperparameter that may depend on the round t (i.e., sample size) and we also define $\hat{\beta}_0 = (0, \ldots, 0)$ for $t = 0$. We call $\hat{\beta}_t$ the LASSO estimator after the reward in round t.

In the evaluation of the gap between the estimator $\hat{\beta}_t$ and true parameter β^*, the compatibility condition defined below plays an essential role, as seen in Lemma 3:

Definition 1 (Compatibility Condition)**.** We say that a positive semi-definite matrix $\Sigma \in \mathbb{R}^{d\times d}$ satisfies the compatibility condition if the following inequality holds for some compatibility constant $\phi > 0$ and some active set $\mathcal{S} \subset [d]$:

$$\frac{V^\top \Sigma V}{\|V\|_{\mathcal{S},1}^2} \geq \frac{\phi^2}{|\mathcal{S}|}\,, \tag{4}$$

for any $V \in \mathbb{R}^d$ such that $\|V\|_{\mathcal{S}^c,1} \leq 3\|V\|_{\mathcal{S},1}$. Here $\mathcal{S}^c := [d]\backslash\mathcal{S}$ and the norm $\|\cdot\|_{\mathcal{T},1}$ for a set $\mathcal{T} \subset [d]$ indicates the L^1-norm for the indices $i \in \mathcal{T}$, i.e., $\|V\|_{\mathcal{T},1} := \sum_{i\in\mathcal{T}} |V_i|$.

In this paper, we fix the active set $\mathcal{S} = \{i \in [d] \mid \beta_i^* \neq 0\}$ and $|\mathcal{S}| = s_0$. We also define a map $\phi_{\mathcal{S}} : \mathbb{R}^{d\times d} \to \mathbb{R}_{\geq 0}$ by

$$\phi_{\mathcal{S}}(\Sigma) := \min_{V\in\mathcal{D}} \sqrt{|\mathcal{S}|\frac{V^\top \Sigma V}{\|V\|_{\mathcal{S},1}^2}}\,, \tag{5}$$

where $\mathcal{D} := \{V \in \mathbb{R}^d \mid \|V\|_{\mathcal{S}^c,1} \leq 3\|V\|_{\mathcal{S},1}\}$. The statement that Σ satisfies the compatibility condition is equivalent to $\phi_{\mathcal{S}}(\Sigma) > 0$. In the regret analysis of the sparse linear bandit, the empirical and expected Gram-matrix, given by the first and second definitions below respectively, are subject to the compatibility condition:

$$G_t := \frac{1}{t}\sum_{s=1}^{t} X_{a_s}^s X_{a_s}^{s\top}\,,\;\; \bar{G}_t := \frac{1}{t}\sum_{s=1}^{t} \mathbb{E}[X_{a_s}^s X_{a_s}^{s\top} \mid \mathcal{F}'_{s-1}]\,. \tag{6}$$

Here we also define $\phi_t := \phi_{\mathcal{S}}(G_t)$ and $\bar{\phi} := \min_{t\in[T]} \phi_{\mathcal{S}}(\bar{G}_t)$ for convenience. We note that $\bar{\phi}$ depends on both the arm feature distribution and the arm selection policy. To clarify the dependence of arm feature distribution $P(\mathcal{X})$, we sometimes write the expected Gram-matrix as $\bar{G}_t(P)$.

Under Assumption 1, it is known that the following two inequalities hold with high probabilities. [1]

Lemma 1 (Lemma 4 in Oh, Iyengar, and Zeevi (2021))**.** *Under Assumption 1, for any $\delta > 0$, the following inequality holds:*

$$\frac{1}{t}\left\|\sum_{s=1}^{t} \epsilon_s X_{a_s}^s\right\|_\infty \leq x_{\max}\sigma\sqrt{\frac{\delta^2 + 2\log d}{t}}\,, \tag{7}$$

with probability at least $1 - e^{-\delta^2/2}$. Here ϵ_s is the σ-sub-Gaussian noise in Eq. (1).

Lemma 2 (Corollary 2 in Oh, Iyengar, and Zeevi (2021))**.** *Under Assumption 1, for $t \geq T_0 := \log(d(d-1))/\kappa(\bar{\phi})$ where $\kappa(\bar{\phi}) := \min(2 - \sqrt{2}, \bar{\phi}^2/(256 x_{\max}^2 s_0))$, the following inequality holds:*

$$\phi_t^2 \geq \frac{\bar{\phi}^2}{2}\,. \tag{8}$$

with probability at least $1 - e^{-t\kappa(\bar{\phi})^2}$.

When the above two highly probable events hold, and if $\bar{\phi} > 0$, the gap between the true parameters β^* and the estimator $\hat{\beta}_t$ can be bounded by the inverse of $\bar{\phi}$, as stated in the following lemma.

Lemma 3. *If the inequalities* (7) *and* (8) *hold, and $\bar{\phi} > 0$, then by setting $\lambda_t \geq 4x_{\max}\sigma\sqrt{(\delta^2 + 2\log d)/t}$, we have*

$$\|\beta^* - \hat{\beta}_t\|_1 \leq \frac{8\lambda_t s_0}{\bar{\phi}^2}\,. \tag{9}$$

3.3 The Compatibility Constant in the Greedy Algorithm

We here present the contribution of the compatibility constant $\bar{\phi}$ in the greedy algorithm. The greedy algorithm chooses the arm a_t that satisfies $a_t = \mathrm{argmax}_{k\in[K]} X_k^{t\top}\hat{\beta}_{t-1}$ for each round t, where $\hat{\beta}_{t-1}$ is the LASSO estimator defined in Eq. (3). Under the greedy policy, we obtain the following regret bound from Lemmas 1, 2, and 3:

Lemma 4. *Under Assumption 1, if $\bar{\phi} > 0$, the expected regret for the greedy algorithm is upper bounded by:*

$$R(T) \leq 2x_{\max}b\left(\frac{1 + \log(d(d-1))}{\kappa(\bar{\phi})^2} + \frac{\pi^2}{3}\right) + \frac{128 s_0 x_{\max}^2\sigma}{\bar{\phi}^2}\sqrt{(4\log T + 2\log d)T}\,, \tag{10}$$

where $\kappa(\bar{\phi}) := \min(2 - \sqrt{2}, \bar{\phi}^2/(256x_{\max}^2 s_0))$.

It follows from Eq. (10) that $R(T) = O(\frac{1}{\bar{\phi}^2}\sqrt{(\log T + \log d)T})$ if $\bar{\phi}$ is non-zero. We note that in this regret analysis, no assumptions other than Assumption 1 are made for the arm feature distribution.

In the following sections, we will discuss what arm feature distributions satisfy the condition $\bar{\phi} > 0$ in Lemma 4. Unfortunately, $\bar{\phi} > 0$ does not hold for all arm feature distributions. In the existing works, this has been shown in a restricted arm feature distribution where an approximate origin-symmetric condition is satisfied, as we will introduce in the next section. Our goal is to show that wider classes of the arm feature distribution satisfy $\bar{\phi} > 0$, and are *greedy-applicable* in the sense that they have a regret upper bound of Eq. (10) for the greedy algorithm.

3.4 Existing Assumptions for the Arm Feature Distribution

In this section, we present three assumptions employed in the existing studies for $\bar{\phi} > 0$, noting in particular that the relaxed symmetry (Assumption 3) severely restricts the arm feature distribution.

The first assumption states that the arm feature distribution must have some diversity:

[1] We give all the proofs in Ichikawa et al. (2023).

Assumption 2. *The expected Gram-matrix with random arm selection satisfies the compatibility condition for a compatibility constant $\phi_0 > 0$:*

$$\phi_{\mathcal{S}}\left(\frac{1}{K}\sum_{k=1}^{K}\mathbb{E}\left[X_k X_k^\top\right]\right) > \phi_0 \tag{11}$$

This assumption is commonly employed in sparse linear bandit algorithms, including the greedy algorithm, and does not strongly constrain the arm feature distribution. We note, however, that this assumption alone does not guarantee sample diversity (i.e., $\bar{\phi} > 0$) under the greedy arm selection policy.

Secondly, the following approximate origin-symmetric condition is imposed for the arm feature distribution $P(X_1,\dots,X_K)$ supported by $\mathrm{supp}(P) \subset (\mathbb{R}^d)^K$ (Oh, Iyengar, and Zeevi 2021; Ariu, Abe, and Proutière 2022):

Assumption 3 (Relaxed Symmetry (RS))**.** *If $\{X_1,\dots,X_K\} \in \mathrm{supp}(P)$, then $\{-X_1,\dots,-X_K\} \in \mathrm{supp}(P)$. Moreover, there exists $1 \le \nu < \infty$ that satisfies $P(-X_1,\dots,-X_K)/P(X_1,\dots,X_K) \le \nu$ for any $\{X_1,\dots,X_K\} \in \mathrm{supp}(P)$.*

We stress that distributions with relaxed symmetry are limited. In particular, it cannot be applied to cases with origin-asymmetric supports. For example, this assumption cannot include arm features that only take positive values or are represented by origin-asymmetric discrete values. One of the main motivations for our research is the relaxation of this origin-symmetric assumption.

Remark 1. *One might think that the arm feature distribution could be made to satisfy the origin-symmetric support by transforming the features, for example, by constant shift or scale transformation. However, when two or more axes of a feature distribution have asymmetric support that is correlated, feature engineering generally cannot guarantee the RS condition.*

A simple case is when labels have a hierarchical structure. For example, if two binary labels show mammal and dog, respectively, (0, 1) meaning "not mammal but dog" is not possible. Transformations that mix axes cannot be performed because they break the sparse structure, and constant shifts and scale transformations for each axis cannot provide origin symmetric support for the discrete distribution that takes values of (0,0), (1,0), (1,1). Such a label structure is common in the recommender system, where the category of items often has a hierarchical structure (e.g., books →academic books →computer science).

While in the case of $K = 2$, the above assumptions ensure $\bar{\phi} > 0$ for the greedy algorithm, the following third assumption is further required in the case of $K > 2$:

Assumption 4 (Balanced Covariance)**.** *For any $i \in \{2,\dots,K-1\}$, any permutation $\pi : [K] \to [K]$, and any fixed $\beta \in \mathbb{R}^d$, there exists a constant $C_{\mathrm{BC}} < \infty$ that satisfies:*

$$\begin{aligned}&\mathbb{E}\left[X_{\pi(i)}X_{\pi(i)}^\top I[X_{\pi(1)}^\top\beta \le \cdots \le X_{\pi(K)}^\top\beta]\right]\\ &\quad\preceq C_{\mathrm{BC}}\mathbb{E}\Big[(X_{\pi(1)}X_{\pi(1)}^\top + X_{\pi(K)}X_{\pi(K)}^\top)\\ &\qquad\qquad I[X_{\pi(1)}^\top\beta \le \cdots \le X_{\pi(K)}^\top\beta]\Big]\,. \end{aligned} \tag{12}$$

Oh, Iyengar, and Zeevi (2021) has shown that if the arm features are generated i.i.d., the coefficients C_{BC} are of finite value, but of the exponential order of K for general distributions. They conjecture that the coefficients would not be so large from the observation of the experimental results. We give the proof of $\bar{\phi} > 0$ under Assumptions 2, 3 and 4 in Ichikawa et al. (2023).

4 Investigation of the Greedy-Applicable Distributions

In this section, we show that the applicability of the greedy algorithm can be extended to wider classes of arm feature distributions by proposing the following two aspects: 1) distributions having a mixture component of a greedy-applicable distribution are also greedy-applicable (Theorem 1), and 2) several representational function classes are greedy-applicable distributions (Section 4.1 and 4.2).

Remark 2. *As noted at the end of Section 3.3, the regret analysis of Lemma 4 does not require assumptions about the arm feature distribution other than Assumption 1. Assumptions 2, 3 and 4 are used solely to show that $\bar{\phi} > 0$. Therefore, Assumptions 2, 3 and 4 can be replaced by other assumptions that lead to $\bar{\phi} > 0$ without changing the regret upper bound in Lemma 4.*

We note that while our analysis focuses on the minimum value of the compatibility constant, $\phi_{\mathcal{S}}$, it can easily be replaced by operators such as minimum eigenvalue, or restricted minimum eigenvalue for the matrix, which also measure the diversity of a matrix.

We conduct our analysis under the following assumption for the arm selection policy: [2]

Assumption 5. *Under given arm features $\mathcal{X}^t = \{X_1^t,\dots,X_K^t\} \in (\mathbb{R}^d)^K$ at round t, the arm selection probability for arm i is described by the greedy policy:*

$$\begin{aligned}&P(\textit{Select } i \mid \mathcal{X}^t, \beta_{t-1})\\ &\quad := \frac{\prod_{j\neq i} I\left[\beta_{t-1}^\top X_i^t \ge \beta_{t-1}^\top X_j^t\right]}{\sum_{i'=1}^{K}\prod_{j\neq i'} I\left[\beta_{t-1}^\top X_{i'}^t \ge \beta_{t-1}^\top X_j^t\right]}\,, \end{aligned} \tag{13}$$

where the parameter $\beta_{t-1} \in \mathbb{R}^d$ is determined from information prior to $\mathcal{X}^t$ and r^t. The denominator is for random selection when tying occurs.

[2] We note that the assumption for the policy can be more general: $P(\text{Select } i \mid \mathcal{X}^t, \mathcal{B}^{t-1}) = f_i(\beta_{1,t-1}^\top X_1^t, \dots, \beta_{K,t-1}^\top X_K^t)$. where $\mathcal{B} := \{\beta_{1,t-1},\dots\beta_{K,t-1}\} \in \Theta \subset (\mathbb{R}^d)^K$. The subsequent theorems are easy to extend, and the greedy policy is considered here for the sake of clarity.

We define the class of arm feature distributions for which sample diversity is guaranteed under Assumption 5 as follows:

Definition 2. The arm feature distribution $P(\mathcal{X})$ is a ϕ_0-*greedy-applicable distribution* if, under Assumption 5, there exists $\phi_0 > 0$ such that the expected Gram-matrix $\bar{G}_t$ satisfies $\phi_{\mathcal{S}}(\bar{G}_t) > \phi_0$.

Then, the following key property holds for the positivity of the compatibility constant:

Theorem 1. *If an arm feature distribution $P(\mathcal{X})$ is a mixture of a PDF $Q(\mathcal{X})$ and a ϕ_0-greedy-applicable distribution $\tilde{P}(\mathcal{X})$, i.e., $P(\mathcal{X}) = c\tilde{P}(\mathcal{X}) + (1-c)Q(\mathcal{X})$ for a constant $0 < c < 1$, then $P(\mathcal{X})$ is a $c\phi_0$-greedy-applicable distribution.*

The theorem indicates that the proof that a class of PDF is greedy-applicable gives a theoretical guarantee to a very wide range of distributions that have this class as their mixture component.[3] Currently, the only general greedy-applicable distributions are the ones introduced in the previous section (i.e., distributions that satisfy Assumption 2, 3, and 4). Below, we propose several new greedy-applicable classes and show the wide applicability of the greedy algorithm.

4.1 Basic Assumptions for the Arm Feature Distribution

Here, we introduce two key assumptions for our analysis: 1) at least one arm distribution is independent of the others, and 2) such an arm will be selected by the algorithm with a positive probability. Intuitively, choosing this independent arm contributes to exploring the arm features if it generates diverse samples. In other words, these assumptions simplify the analysis of the greedy algorithm's applicability by reducing it to a discussion on the feature distribution of the single independent arm, which will be discussed in Section 4.2.

Formally, these two assumptions are described as follows:

Assumption 6. *There exists at least one arm $i \in [K]$ that is independent of the other arms: $P(\mathcal{X}) = P(\mathcal{X}\backslash\{X_i\})P_i(X_i)$.*

Assumption 7. *For i defined in Assumption 6, the marginalized arm selection probability has a positive lower bound with respect to $\beta \in \mathbb{R}^d$ under the greedy policy given in Assumption 5: $\inf_{\beta\in\mathbb{R}^d} P(\text{Select } i \mid \beta) > 0$.*

The second assumption is made to avoid the possibility that arm i is never selected under a certain β. If β satisfying $P(\text{Select } i \mid \beta) = 0$ exists, then constraining P_i alone cannot guarantee the sample diversity, as in the worst case, any $X \sim P_i(X)$ will not be sampled.

To clarify the requirements of the assumptions, we give two application examples below:

Example 1 The simplest example satisfying Assumptions 6 and 7 is when all arm features are generated independently from the same distribution. For instance, consider a case where a set of recommendation candidates is given each week, and the recommendation system selects an item from that set and measures the click-through rate. Suppose each candidate is selected independently and uniformly at random from all items. If we assume that the expectation of the click-through rate is linear with respect to the item features, maximizing the cumulative click-through rate can be related to our bandit problem. In this scenario, it is obvious that Assumption 6 holds and, since all candidates are selected by the same probability, Assumption 7 also holds.

Example 2 Another example is the case where all arm features are generated independently but from different distributions. Consider a book recommendation that has categorical tags of science and fiction. We consider two arms (i.e., two recommendation candidates), where the first and second candidates are selected from independent uniform distributions of science books and fiction books, respectively. Suppose the set of the fiction books includes sci-fi books with both science and fiction tags. In this scenario, Assumption 6 again obviously holds. Moreover, since the fiction books include the sci-fi books with the science tag, the candidate of the 'fiction' arm can be selected even if the greedy algorithm heavily favors the science tag. Therefore, Assumption 7 also holds for the 'fiction' arm.

Remark 3. *We note the relation of our assumptions to the existing assumptions: Assumption 3 allows correlation between all arms and does not request Assumption 7. However, as mentioned previously, Assumption 3 does not allow distributions with asymmetric support. Also, for non-i.i.d. arms, it is necessary to assume Assumption 4. On the other hand, Assumptions 6 and 7 enable us to discuss the greedy-applicability of the arm feature distributions that cannot be handled by the conventional analysis. We also note that it is sufficient for the arm feature distribution to satisfy either the conventional assumptions or those we propose. This condition can be further relaxed by Theorem 1 to the claim that such a distribution is only required to be a component of a mixture.*

As we have mentioned at the beginning of this section, we will focus on the feature distribution of the single independent arm in the next section. We define the distribution $P_i(X)$ that ensures the greedy-applicability of $P(\mathcal{X})$ in the following term:

Definition 3. We call $P_i(X)$ the *basis of the greedy-applicable distribution*, if there exists a positive constant $\phi_0 > 0$ and $P(\mathcal{X})$ is a ϕ_0-greedy-applicable distribution under Assumptions 5, 6, and 7.

In the next section, we provide several bases, omitting index i in $P_i(X_i)$ for brevity.

4.2 Proposal for Several Bases

For the sake of our analysis, we first define the following time-independent expected Gram-matrix :

$$\tilde{G}_\beta := \sum_k \int X_k X_k^\top P(\text{Select } k \mid \mathcal{X}, \beta) P(\mathcal{X}) \prod_{k'=1}^{K} dX_{k'} \,. \tag{14}$$

[3]The coefficient for specific $\tilde{P}$s and situations where $\tilde{P}$ approximates P are presented in Ichikawa et al. (2023).

In this section, the analysis is performed for $\tilde{G}_\beta$ instead of $\bar{G}_t$, according to the following lemma:

Lemma 5. *Under Assumption 5, if $\phi_{\mathcal{S}}(\tilde{G}_\beta) \geq \phi_0$ for any $\beta \in \mathbb{R}^d$, then $\phi_{\mathcal{S}}(\bar{G}_t) \geq \phi_0$.*

Gaussian Mixture Basis First, we propose a basis where $P(X)$ can be decomposed by a sum of finite Gaussian distributions.

Definition 4. Gaussian mixture basis $P_{GM}(X)$ is a PDF that can be decomposed by a sum of finite Gaussian distribution: $P_{GM}(X) = \sum_{n=1}^N w_n \mathcal{N}(X \mid \mu_n, \Sigma_n)$, where $\mu_n \in \mathbb{R}^d$ is a mean vector and $\Sigma_n \in \mathbb{R}^{d\times d}$ is a positive definite covariance matrix for each Gaussian distribution. The weight $w_n > 0$ satisfies $\sum_n w_n = 1$.

Then, the lower bound for $\tilde{G}_\beta$ is given by the following theorem:

Lemma 6. *Under Assumption 5, 6, and 7, if $P_i(X)$ is the Gaussian mixture basis $P_{GM}(X)$, then the following lower bound holds:*

$$\tilde{G}_\beta \succeq \sum_{n=1}^N w_n c_n(\beta)(\Sigma_n + \mu_n\mu_n^\top)\,, \tag{15}$$

where $c_n(\beta) > 0$ is a β-dependent positive constant.

Since Σ is a positive definite matrix, the following is obvious from Lemmas 5 and 6:

Theorem 2. *$P_{GM}(X)$ is a basis of the greedy-applicable distribution.*

Low-rank Gaussian Mixture and Discrete Basis Lemma 6 shows that $\tilde{G}_\beta$ can be bounded by a weighted sum of the second moments of each Gaussian component. Importantly, if we allow $c_n(\beta) = 0$, this lower bound also holds in the limit where the eigenvalues of Σ_n are taken to be zero. We provide a useful lemma for the coefficient $c_n(\beta)$:

Lemma 7. *$c_n(\beta) = 0$ if and only if the following two conditions hold:*

$$\beta^\top \Sigma_n \beta = 0\,, \textit{and } P(\textit{Select } i \mid X_i = \mu_n, \beta) = 0\,, \tag{16}$$

where $P(\textit{Select } i \mid X_i, \beta)$ is the marginalized probability distribution for $\{X_j \mid j \in [K]\backslash\{i\}\}$.

Corollary 2.1. $\lim_{\beta^\top\Sigma_n\beta\to 0} c_n(\beta) = P(\textit{Select } i \mid X_i = \mu_n, \beta)$.

The limit operation allows us to include the Gaussian mixture distribution with a low-rank covariance matrix and discrete distribution in the theory. For a positive semi-definite matrix $\Sigma^{d'} \in \mathbb{R}^{d\times d}$ with rank $d' \leq d$ and diagonalized by an orthogonal matrix $R \in \mathbb{R}^{d\times d}$ as $R^\top \Sigma^{d'} R = \mathrm{diag}(\lambda_1, \ldots, \lambda_{d'}, 0, \ldots, 0)$ where $\lambda_i > 0$ for $i \in [d']$, we define the low-rank Gaussian distribution as follows:

$$\tilde{\mathcal{N}}(X \mid \mu, \Sigma^{d'}) := \frac{\prod_{j=d'+1}^{d} \delta((RX - R\mu)_j)}{(2\pi)^{d'/2}(\prod_{i=1}^{d'} \lambda_i)^{1/2}} e^{-\frac{1}{2}(X-\mu)^T \Lambda^{d_n}(X-\mu)}\,, \tag{17}$$

where $\Lambda^{d'} := R\,\mathrm{diag}(\lambda_1^{-1}, \ldots, \lambda_{d'}^{-1}, 0, \ldots, 0)R^\top$ and $\delta(x)$ is the delta function. We define the low-rank Gaussian mixture basis as follows:

Definition 5. Low-rank Gaussian mixture basis $P_{LGM}(X)$ is a PDF that can be decomposed as a sum of finite low-rank Gaussian distribution: $P_{LGM}(X) = \sum_{n=1}^N w_n \tilde{\mathcal{N}}(X \mid \mu_n, \Sigma_n^{d_n})$. Moreover, P_{LGM} satisfies the following condition for a positive constant $\phi_0 > 0$:

$$\inf_{\beta\in\mathbb{R}^d} \phi_{\mathcal{S}}\left(\sum_n c_n(\beta) w_n (\Sigma_n^{d_n} + \mu_n\mu_n^\top)\right) \geq \phi_0\,. \tag{18}$$

Then, we obtain the following theorem:

Theorem 3. *$P_{LGM}(X)$ is a basis of the greedy-applicable distribution.*

We note that not all Gaussian mixtures with low-rank covariance matrix are included in the class, and that the condition indicated in Eq. (18) is necessary for their second moment.

Trivially, the limit operation $\Sigma_n \to 0$ can represent a discrete distribution.

Definition 6. Discrete basis $P_D(X)$ is a PDF that can be described by a discrete probability distribution: $P_D(X = \mu_n) = p_n$, where $p_n > 0$ satisfies $\sum_n p_n = 1$, and $\mu_n \in \mathbb{R}^d$ is an element of a set $\mathcal{M} := \{\mu_1, \ldots, \mu_N\}$. Moreover, P_D satisfies the following condition for a positive constant $\phi_0 > 0$:

$$\inf_{\beta\in\mathbb{R}^d} \phi_{\mathcal{S}}\left(\sum_n c_n(\beta) p_n \mu_n\mu_n^\top\right) \geq \phi_0\,. \tag{19}$$

Corollary 3.1. *$P_D(X)$ is a basis of the greedy-applicable distribution.*

Radial Basis Due to the nature of the Gaussian distribution, P_{GM} and P_{LGM} cannot include distributions with truncation. As a basis for the truncated distribution, we consider PDF of radial function.

Definition 7. Radial mixture basis $P_R(X)$ is a PDF that can be decomposed by a sum of finite radial distributions: $P_R(X) = \sum_{n=1}^N w_n Q_n(X \mid \mu_n)$, where $\mu_n \in \mathbb{R}^d$ is a vector and PDF $Q_n(X \mid \mu_n)$ is a radial function: $Q_n(X \mid \mu_n) = f_n(\|X - \mu_n\|_2)$ that satisfies $\int f_n(\|X\|_2)dX = 1$.

The radial function f_n can include, for example, truncated uniform distribution and truncated standard normal distribution. Then, the following theorem holds:

Theorem 4. *$P_R(X \mid \mu)$ is a basis of the greedy-applicable distribution.*

We summarize the formulae for the $\phi_{\mathcal{S}}(\bar{G}_t)$ lower bound of each basis in Ichikawa et al. (2023).

We here mention the relationship between our analysis and the smoothed analysis (Kannan et al. 2018; Sivakumar, Wu, and Banerjee 2020). In the smoothed analysis, it is assumed that the arm features are generated adversarially, and then observed with stochastic perturbations from a truncated isotropic normal distribution. Sivakumar, Wu, and Banerjee (2020) showed that the greedy algorithm in the sparse linear bandit can achieve an $O(\sqrt{sdT})$ regret upper bound in the perturbed adversarial setting. The radial basis $P_R(X) = \sum_{n=1}^N w_n Q_n(X \mid \mu_n)$ is considered to correspond to a specific stochastic setting of the smoothed analysis. That is, we can regard the radial basis as a distribution where the arm feature μ_n is chosen with probability w_n and then perturbed by $Q_n(X \mid \mu_n)$. In our analysis, we include general radial distributions other than the truncated isotropic normal distribution and also examine the applicability of the greedy algorithm for P_{GM}, P_{LGM}, and P_D. While we expect these bases to retain their properties even in the perturbed adversarial setting, we leave a more detailed analysis for future work.

4.3 Examples

In this section, we demonstrate how the theorems of the previous section apply and show the greedy-applicability to specific examples that could not be dealt with in the analysis of previous studies. In addition, we also performed a numerical experiment with artificial data to empirically validate our claim, which is given in Ichikawa et al. (2023). Below, we use the following lemma.

Lemma 8. *Let us define positive semi-definite matrices $\Lambda_n \in \mathbb{R}^{d\times d}$ and positive coefficients $w_n > 0$, $w'_n > 0$ for $n \in [N]$. If $\phi_{\mathcal{S}}(\sum_{n=1}^N w_n\Lambda_n) > 0$, then $\phi_{\mathcal{S}}(\sum_{n=1}^N w'_n\Lambda_n) > 0$.*

Consider that $d = 2$ and the arms have two independent binary features (x_1, x_2) as features and is realized in arm i with the following non-zero probabilities: p_1 for $\mu_1 = (0,0)$, p_2 for $\mu_2 = (1,0)$, p_3 for $\mu_3 = (0,1)$, and p_4 for $\mu_4 = (1,1)$, where $p_1+p_2+p_3+p_4 = 1$. We also assume that arm i is independent of the other arms (for Assumption 6) and $\inf_\beta P(\text{Select } i \mid X_i, \beta) > 0$ under the realization of $X_i = \mu_2, \mu_3$, or μ_4 in the greedy algorithm (for Assumption 7, Lemma 7, and Corollary 2.1). Since X_i does not take negative values, it is clearly a distribution for which Assumption 3 is not valid. In our analysis, from Lemma 6, we obtain $\tilde{G}_\beta \succeq \sum_{n=1}^4 c_n(\beta)p_n\mu_n\mu_n^\top \succeq \sum_{n=2}^4 c_n(\beta)p_n\mu_n\mu_n^\top$, and from Corollary 2.1, $\inf_\beta c_n(\beta) > 0$ for $n = 2, 3$ and 4. Meanwhile,

$$\sum_{n=2}^4 p_n\mu_n\mu_n^\top = \begin{pmatrix} p_2+p_4 & p_4 \\ p_4 & p_3+p_4 \end{pmatrix} \tag{20}$$

is a positive definite matrix. Then, Lemma 8 derives that there exists $\phi_0 > 0$ such that $\inf_\beta \phi_{\mathcal{S}}(\sum_{n=2}^4 c_n(\beta)p_n\mu_n\mu_n^\top) > \phi_0$, which indicates that the distribution of this example is ϕ_0-greedy applicable. We note that $p_3 = 0$ is the case shown in Remark 1 that cannot satisfy Assumption 3 even with the constant shift.

As a next example, consider $d = 2$, two independent arms, and both features uniformly distributed within the region $\{(x_1, x_2) \mid 0 \le x_1 \le 1, 0 \le x_2 \le 1, \sqrt{x_1^2 + x_2^2} \ge 0.1\}$. Again, this arm feature distribution does not satisfy Assumption 3 even with the constant shift. In our analysis, we see that Assumptions 6 and 7 are satisfied. Moreover, the distribution of both arms is a mixture of the radial basis $Q(X) = f(\|X - (1/2, 1/2)\|_2)$ and the remaining, where $f(r) : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies $f(r) = 4/\pi$ for $0 \le r \le 1/2$ and 0 otherwise. Therefore, from Theorems 1 and 4, this arm feature distribution is a (some) ϕ_0-greedy applicable distribution.

5 Application to Several Algorithms

Our analysis can be applied to many existing algorithms. We illustrate some of them in this section. While the application examples in this section focus on the sparse settings, an application to the dense parameter setting is also shown in Ichikawa et al. (2023).

Greedy algorithm (Oh, Iyengar, and Zeevi 2021) In the analysis for the greedy algorithm, Assumptions 2, 3 and 4 are imposed on the arm feature distribution to guarantee $\bar{\phi} > 0$ (Lemma 10 of Oh, Iyengar, and Zeevi (2021)). Our assumptions can replace these assumptions: Under Assumptions 6 and 7, and P_i in Assumption 6 being one of the bases proposed in Section 4.2, we can conclude $\bar{\phi} > 0$. Also, as Theorem 1 shows, if the arm feature distribution has a mixture component of a greedy method-applicable distribution, then $\bar{\phi} > 0$.

Thresholded lasso bandit (Ariu, Abe, and Proutière 2022) The thresholded lasso bandit estimates the support of β^* each round and a greedy arm selection policy is performed by the inner product of the arm features and the estimated parameter for β^* on this support. The relaxed symmetry and the balanced covariance are introduced to ensure proper support estimation under the greedy policy. Specifically, these assumptions are again used for Lemma 10 of Oh, Iyengar, and Zeevi (2021) in Lemma 5.4. Therefore, as with the greedy algorithm, our proposed classes can be used as the assumptions of the arm feature distribution.

Greedy algorithm for the combinatorial setting Our analysis is applicable to the combinatorial bandit setting, where no regret upper bound for the sparsity-agnostic greedy algorithm is still given. Here we consider the setting where in each round, L of the K arms are selected and their respective rewards are observed. Suppose that in each round t, the selection policy determines a set of arms $\mathcal{I}_t \subset [K]$ where $|\mathcal{I}_t| = L$ to be selected. The reward under selection criteria is given by: $r_t := \sum_{a_t\in\mathcal{I}_t} X_{a_t}^{t\top}\beta^* + \epsilon_t$, whereas the optimal reward r_t^* is given under the optimal arm-set selection $\mathcal{I}_t^*$. The empirical and expected Gram-matrices are given by:

$$G_t := \frac{1}{Lt}\sum_{s=1}^t \sum_{a_s\in\mathcal{I}_s} X_{a_s}^s X_{a_s}^{s\top},$$
$$\bar{G}_t := \frac{1}{Lt}\sum_{s=1}^t \mathbb{E}\left[\sum_{a_s\in\mathcal{I}_s} X_{a_s}^s X_{a_s}^{s\top} \mid \mathcal{F}'_{s-1}\right], \tag{21}$$

respectively. Then, Lemma 1, 2 and 3 do not depend on the arm selection policy and hold for the combinatorial setting by simply replacing t with Lt.

The greedy policy for the combinatorial setting is to choose the top-L arms:

Assumption 8. *Under given arm features* $\mathcal{X}^t = \{X_1^t, \ldots, X_K^t\} \in (\mathbb{R}^d)^K$ *at round* t*, the arm selection probability for a set of arms* $\mathcal{I}_t$ *is given by:* $P(\text{Select}\, \mathcal{I}_t \mid \mathcal{X}^t, \beta_{t-1}) \propto \prod_{i \in \mathcal{I}_t} \prod_{j \notin \mathcal{I}_t} I(\beta_{t-1}^\top X_i^t \geq \beta_{t-1}^\top X_j^t)$*, where the parameter* $\beta_{t-1} \in \mathbb{R}^d$ *is determined from information prior to* $\mathcal{X}^t$ *and* r^t.

Then, similar to Lemma 4, the following statement holds:

Corollary 4.1. *Under Assumption 1, and if* $\min_{t \in [T]} \phi_{\mathcal{S}}(\bar{G}_t) > \bar{\phi}$*, the expected regret for the greedy algorithm is upper bounded by:* $R(T) = O\left(\frac{1}{\bar{\phi}^2}\sqrt{(\log(dT)T}\right)$.

The positivity of $\bar{\phi}$ is therefore also important under the combinatorial setting. In this regard, the following theorem holds.

Theorem 5. *Under Assumption 6, 7, and 8, and if* $P_i(X_i)$ *is described by* P_{GM}*,* P_{LGM}*,* P_D*, or* P_R*, then there exists* $\phi_0' > 0$ *that satisfies* $\phi_{\mathcal{S}}(\bar{G}_t) > \phi_0'$.

6 Discussion and Conclusion

In this paper, the applicability of the greedy algorithm in the sparse linear contextual bandits was considered. We showed that the regret upper bound of the greedy algorithm is guaranteed over a wider range of the arm feature distribution than that covered by the previous works. In addition, we demonstrated that our analytical approach is not restricted to the simple greedy algorithm, but can be applied in a variety of settings.

On the other hand, the analysis still leaves much room for development. First, the size of ϕ_0 has not been fully explored, in particular how the shape of the distribution, the dimension of the arm features and the number of arms affect it.[4] Secondly, for truncated distributions, although we gave the radial basis, it is conceivable that a more representational basis could exist. Finally, we assumed that at least one arm is independent of the other arms, as stated in Assumption 6. Although this is looser than the assumption of all arms being independent, which is assumed in many analyses, it is an interesting subject for future research on what class of arm feature distributions can be greedy-applicable when all arms are correlated.

Acknowledgments

HS was supported by JSPS KAKENHI Grant Numbers JP17K12646, JP21K17708, and JP21H03397, Japan. TF was supported by JSPS KAKENHI Grant Numbers JP20H05965, JP21K11759, and JP21H03397, Japan. NK was supported by JSPS KAKENHI Grant Numbers JP22H05001, JP20H05795, and JP21H03397, Japan and JST ERATO Grant Number JPMJER2301, Japan. KK was supported by JSPS KAKENHI Grant Numbers JP22H05001 and JP20A402, Japan.

[4]The quantitative lower bound discussion is difficult without further specification of the arm feature distribution. Specifically, $P(\text{Select } i \mid X_i, \beta)$, which appears in the lower bound evaluation, has almost no constraints on its shape under our assumptions. Avoiding additional assumptions, we give formulae that can, in principle, compute a lower bound for $\bar{\phi}$ under a given arm feature distribution, which is summarized in Ichikawa et al. (2023).

References

Abbasi-Yadkori, Y.; Pál, D.; and Szepesvári, C. 2011. Improved Algorithms for Linear Stochastic Bandits. In *NIPS 2011*, 2312–2320.

Abbasi-Yadkori, Y.; Pál, D.; and Szepesvári, C. 2012. Online-to-Confidence-Set Conversions and Application to Sparse Stochastic Bandits. In *AISTATS 2012*, volume 22, 1–9.

Ariu, K.; Abe, K.; and Proutière, A. 2022. Thresholded Lasso Bandit. In *ICML 2022*, volume 162, 878–928.

Auer, P. 2002. Using Confidence Bounds for Exploitation-Exploration Trade-offs. *JMLR*, 3: 397–422.

Bastani, H.; and Bayati, M. 2020. Online Decision Making with High-Dimensional Covariates. *Oper. Res.*, 68(1): 276–294.

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

Bouneffouf, D.; and Rish, I. 2019. A Survey on Practical Applications of Multi-Armed and Contextual Bandits. *arXiv preprint arXiv:1904.10040*.

Carpentier, A.; and Munos, R. 2012. Bandit Theory meets Compressed Sensing for high dimensional Stochastic Linear Bandit. In *AISTATS 2012*, volume 22, 190–198.

Chu, W.; Li, L.; Reyzin, L.; and Schapire, R. E. 2011. Contextual Bandits with Linear Payoff Functions. In *AISTATS 2011*, volume 15, 208–214.

Dani, V.; Hayes, T. P.; and Kakade, S. M. 2008. Stochastic Linear Optimization under Bandit Feedback. In *COLT 2008*, 355–366.

Durand, A.; Achilleos, C.; Iacovides, D.; Strati, K.; Mitsis, G. D.; and Pineau, J. 2018. Contextual Bandits for Adapting Treatment in a Mouse Model of de Novo Carcinogenesis. In *MLHC 2018*, volume 85, 67–82.

Ichikawa, K.; Ito, S.; Hatano, D.; Sumita, H.; Fukunaga, T.; Kakimura, N.; and Kawarabayashi, K. 2023. New classes of the greedy-applicable arm feature distributions in the sparse linear bandit problem. *arXiv preprint arXiv:2312.12400*.

Kannan, S.; Morgenstern, J.; Roth, A.; Waggoner, B.; and Wu, Z. S. 2018. A Smoothed Analysis of the Greedy Algorithm for the Linear Contextual Bandit Problem. In *NeurIPS 2018*, 2231–2241.

Kim, G.; and Paik, M. C. 2019. Doubly-Robust Lasso Bandit. In *NeurIPS 2019*, 5869–5879.

Lattimore, T.; and Szepesvári, C. 2020. *Bandit algorithms*. Cambridge University Press.

Li, L.; Chu, W.; Langford, J.; and Schapire, R. E. 2010. A contextual-bandit approach to personalized news article recommendation. In *WWW 2010*, 661–670.

Oh, M.; Iyengar, G.; and Zeevi, A. 2021. Sparsity-Agnostic Lasso Bandit. In *ICML 2021*, volume 139, 8271–8280.

Rusmevichientong, P.; and Tsitsiklis, J. N. 2010. Linearly Parameterized Bandits. *Math. Oper. Res.*, 35(2): 395–411.

Sivakumar, V.; Wu, S.; and Banerjee, A. 2020. Structured linear contextual bandits: A sharp and geometric smoothed analysis. In *ICML 2020*, 9026–9035.

Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1): 267–288.

Wang, X.; Wei, M. M.; and Yao, T. 2018. Minimax Concave Penalized Multi-Armed Bandit Model with High-Dimensional Convariates. In *ICML 2018*, volume 80, 5187–5195.