

# MATRIX SKETCHING IN BANDITS: CURRENT PITFALLS AND NEW FRAMEWORK

**Anonymous authors**

Paper under double-blind review

## ABSTRACT

The utilization of sketching techniques has progressively emerged as a pivotal method for enhancing the efficiency of online learning. In linear bandit settings, current sketch-based approaches leverage matrix sketching to reduce the per-round time complexity from  $\Omega(d^2)$  to  $O(d)$ , where  $d$  is the input dimension. Despite this improved efficiency, these approaches encounter critical pitfalls: if the spectral tail of the covariance matrix does not decrease rapidly, it can lead to *linear regret*. In this paper, we revisit the regret analysis and algorithm design concerning approximating the covariance matrix using matrix sketching in linear bandits. We illustrate how inappropriate sketch sizes can result in unbounded spectral loss, thereby causing linear regret. To prevent this issue, we propose Dyadic Block Sketching, an innovative streaming matrix sketching approach that adaptively manages sketch size to constrain global spectral loss. This approach effectively tracks the best rank- $k$  approximation in an online manner, ensuring efficiency when the geometry of the covariance matrix is favorable. Then, we apply the proposed Dyadic Block Sketching to linear bandits and demonstrate that the resulting bandit algorithm can achieve *sublinear regret* without prior knowledge of the covariance matrix, even under the worst case. Our method is a general framework for efficient sketch-based linear bandits, applicable to all existing sketch-based approaches, and offers improved regret bounds accordingly. Additionally, we conduct comprehensive empirical studies using both synthetic and real-world data to validate the accuracy of our theoretical findings and to highlight the effectiveness of our algorithm.

## 1 INTRODUCTION

The Multi-Armed Bandits (MAB) model represents a framework for sequential decision-making under conditions of partial information (Robbins, 1952). In each round, the player selects one of the  $K$  arms to maximize cumulative rewards. The player’s strategy, which guides action choices based on previous observations, is referred to as a policy. The player’s objective is to develop a policy that minimizes regret, which is defined as the difference between the total reward of the optimal policy and that of the chosen policy.

We consider the Stochastic Linear Bandit (SLB) model, a variant of the MAB model under a linear assumption (Abbasi-Yadkori et al., 2011; Auer, 2002; Chu et al., 2011; Dani et al., 2007). In SLB, at round  $t$ , the player selects an arm  $\mathbf{x}_t$  from an alternate set  $\mathcal{X}$ , and then observes the reward  $r_t$ . The expected reward  $E[r_t|\mathbf{x}_t] = \mathbf{x}_t^\top \boldsymbol{\theta}_*$ , where  $\boldsymbol{\theta}_*$  represents unknown coefficients. The player’s objective is to minimize the regret over the total  $T$  rounds, defined as:

$$\text{Regret}_T = \sum_{t=1}^T \max_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\top \boldsymbol{\theta}_* - \sum_{t=1}^T \mathbf{x}_t^\top \boldsymbol{\theta}_*. \quad (1)$$

Utilizing upper confidence bounds and the regularized least squares estimator, Abbasi-Yadkori et al. (2011) introduced the well-known OFUL algorithm, which achieves a regret of  $\tilde{O}(d\sqrt{T})$  and exhibits a computational complexity of  $\Omega(d^2)$ , where  $d$  represents the dimension of the data. However, in real-world decision-making problems, the data dimension  $d$  often increases rapidly, making traditional bandit algorithms excessively time-consuming (Calandriello et al., 2019; Xu et al., 2020; Deshpande & Montanari, 2012; Zhang et al., 2024). To address this issue and eliminate

054 the quadratic dependence on  $d$ , various approaches have adopted streaming sketching methods to  
 055 enhance efficiency. Yu et al. (2017) employ random projection to map high-dimensional arms to  
 056 a low-dimensional subspace. Another line of these works is based on a well-known deterministic  
 057 sketching algorithm—Frequent Directions (FD), which has been proven to offer better theoretical  
 058 guarantees than random projection under the streaming setting (Liberty, 2013; Woodruff et al., 2014;  
 059 Ghashami et al., 2016). Kuzborskij et al. (2019) directly use FD to sketch covariance matrices of  
 060 linear bandits, successfully reducing time complexity from  $\Omega(d^2)$  to  $O(dl + l^2)$  while achieving an  
 061 upper regret bound of  $\tilde{O}\left((1 + \Delta_T)^{\frac{3}{2}}(l + d \log(1 + \Delta_T))\sqrt{T}\right)$ , where  $l < d$  is the sketch size and  
 062  $\Delta_T$  represents the spectral error caused by the shrinking of FD. Building on this foundation, Chen  
 063 et al. (2021) introduce Robust Frequent Directions (RFD) to reduce the order of  $\Delta_T$  and decouple  $d$   
 064 and  $\Delta_T$ , achieving an improved regret bound of  $\tilde{O}\left((\sqrt{l + d \log(1 + \Delta_T)} + \sqrt{\Delta_T})\sqrt{lT}\right)$ .  
 065

066 Despite recent advancements, matrix sketching in linear bandits still faces several pitfalls in practical  
 067 applications (Kuzborskij et al., 2019; Chen et al., 2021; Calandriello et al., 2019). Upon careful  
 068 examination of the previous regret bounds, the  $\sqrt{T}$  term is associated with a spectral error  $\Delta_T$ ,  
 069 which arises from the discrepancy between the sketched and non-sketched regularized least squares  
 070 estimators. The magnitude of the spectral error  $\Delta_T$  relates to the fixed sketch size  $l$  and the spectral  
 071 tail of the covariance matrix, implying that a slow decrease in this tail can contribute to linear regret.  
 072 In practice, this is evident as an inappropriate sketch size can significantly hinder the performance of  
 073 online learning algorithms. Since the spectral information of the covariance matrix is unknown prior  
 074 to online learning, selecting an optimal pre-set, fixed sketch size is challenging. This raises a natural  
 075 question: *Can we adaptively adjust the sketch size in an online manner to avoid the pitfall of linear  
 076 regret in current methods?*

077 In this paper, we demonstrate that the answer is "yes" by developing a novel framework for efficient  
 078 sketch-based linear bandit algorithms. Specifically, this work makes three key contributions:

- 079 • We revisit the fundamental problem of approximating the covariance matrix through matrix  
 080 sketching. We analyze the critical condition for linear regret in sketch-based methods, which  
 081 depends on unpredictable properties of the streaming matrix. From both theoretical and experimental  
 082 perspectives, we demonstrate that the inability of previous methods to avoid the pitfall of linear regret  
 083 stems from the difficulty of pre-setting an appropriate fixed sketch size.
- 084 • We propose Dyadic Block Sketching, a multi-scale matrix sketching method that imposes a  
 085 constraint on the global spectral error by managing the error bound within each block. We prove that  
 086 the cumulative spectral error upper bound from Dyadic Block Sketching conforms to a specified error  
 087  $\epsilon$ . This approach allows the sketch size to be dynamically adjusted to accommodate the given error,  
 088 even without prior knowledge of the matrix structure. Additionally, we demonstrate that Dyadic  
 089 Block Sketching effectively tracks the best rank- $k$  approximation in the streaming setting, aligning  
 090 with the performance of a single deterministic sketch.
- 091 • We introduce an efficient framework for sketch-based linear bandits using Dyadic Block Sketching,  
 092 effectively addressing the pitfall of linear regret in previous works. Our framework is robust, scalable,  
 093 and capable of achieving various regret bounds through different sketching techniques. By tracking  
 094 the best rank- $k$  approximation, our method can significantly reduce the computational cost of linear  
 095 bandits when the covariance matrix has favorable properties.

096  
 097 **Related Work.** Two classes of prior work are particularly relevant to our study: matrix sketching  
 098 algorithms in the unbounded streaming model and sketch-based online learning algorithms. Stream-  
 099 ing matrix sketching methods can be broadly categorized into three groups: The first approach is  
 100 sampling a small subset of matrix rows or columns that approximates the entire matrix (Deshpande  
 101 & Rademacher, 2010; Frieze et al., 2004). The second approach is randomly combining matrix  
 102 rows via random projection. Several results are available in the literature, including random pro-  
 103 jections and hashing (Sarlos, 2006; Achlioptas, 2001). The third approach employs a deterministic  
 104 matrix sketching technique proposed by Liberty (2013), which adapts the well-known MG algo-  
 105 rithm from Misra & Gries (1982) (originally used for approximating item frequencies) to sketch  
 106 a streaming matrix by tracking its frequent directions. For further details, we refer readers to the  
 107 survey by Woodruff et al. (2014). In sketch-based online learning algorithms, most existing work  
 aims to enhance efficiency through sketching. Beyond linear bandits setting, matrix sketching is  
 also employed to accelerate second-order online gradient descent (Luo et al., 2016), online kernel

learning (Calandriello et al., 2017; Luo et al., 2019), stochastic optimization (Gonen et al., 2016), and contextual batched bandits (Zhang et al., 2024).

**Organization.** The remainder of this paper is structured as follows: Section 2 revisits matrix sketching methods in linear bandits and highlights the current pitfalls. Section 3 presents a novel multi-scale sketching method for achieving a constrained global error bound. Section 4 introduces a new framework for efficient sketch-based linear bandits. Section 5 provides a detailed report of the experimental results. Finally, Section 6 concludes the paper and offers a discussion. All proofs and additional algorithmic details are provided in the appendices.

## 2 REVISITING MATRIX SKETCHING IN LINEAR BANDITS

**Notations.** Let  $[n] = \{1, 2, \dots, n\}$ , upper-case bold letters (e.g.,  $\mathbf{A}$ ) represent matrix and lower-case bold letters (e.g.,  $\mathbf{a}$ ) represent vectors. We denote by  $\|\mathbf{A}\|_2$  and  $\|\mathbf{A}\|_F$  the spectral and Frobenius norms of  $\mathbf{A}$ . We define  $|\mathbf{A}|$  and  $\text{Tr}(\mathbf{A})$  as the determinant and trace of matrix  $\mathbf{A}$ . For a positive semi-definite matrix  $\mathbf{A}$ , the matrix norm of vector  $\mathbf{x}$  is defined by  $\|\mathbf{x}\|_{\mathbf{A}} = \sqrt{\mathbf{x}^\top \mathbf{A} \mathbf{x}}$ . For two positive semi-definite matrices  $\mathbf{A}$  and  $\mathbf{B}$ , we use  $\mathbf{A} \succeq \mathbf{B}$  to represent the fact that  $\mathbf{A} - \mathbf{B}$  is positive semi-definite. We use  $\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^\top$  to represent the SVD of  $\mathbf{A}$ , where  $\mathbf{U}$ ,  $\mathbf{V}$  denote the left and right matrices of singular vectors and  $\mathbf{\Sigma} = \text{diag}[\sigma_1, \dots, \sigma_n]$  is the diagonal matrix of singular values in the descending order. We define  $\mathbf{A}_{[k]} = \mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k^\top$  for  $k \leq \text{rank}(\mathbf{A})$  as the best rank- $k$  approximation to  $\mathbf{A}$ , where  $\mathbf{U}_k \in \mathbb{R}^{n \times k}$  and  $\mathbf{V}_k \in \mathbb{R}^{d \times k}$  are the first  $k$  columns of  $\mathbf{U}$  and  $\mathbf{V}$ .

### 2.1 LINEAR BANDITS THROUGH MATRIX SKETCHING

Within the linear bandit setting, the reward for choosing action  $\mathbf{x}_t$  is defined as  $r_t = \mathbf{x}_t^\top \boldsymbol{\theta}_* + z_t$ , where  $\boldsymbol{\theta}_*$  is a fixed, unknown vector of real coefficients, and  $z_t$  denotes a zero-mean random variable. Traditional linear bandit algorithms utilize regularized least squares (RLS) to estimate the unknown weight  $\boldsymbol{\theta}_*$  as

$$\mathbf{A}^{(t)} = \lambda \mathbf{I}_d + \left( \mathbf{X}^{(t)} \right)^\top \mathbf{X}^{(t)} \quad \text{and} \quad \hat{\boldsymbol{\theta}}_t = \left( \mathbf{A}^{(t)} \right)^{-1} \sum_{s=1}^t r_s \mathbf{x}_s, \quad (2)$$

where  $\left( \mathbf{X}^{(t)} \right)^\top = [\mathbf{x}_1^\top, \dots, \mathbf{x}_t^\top]$  is the  $d \times t$  matrix containing all the arms selected up to round  $t$  and  $\lambda$  is the regularization parameter.

Sketch-based linear bandit methods create a smaller matrix (termed sketch matrix)  $\mathbf{S}^{(t)} \in \mathbb{R}^{l \times d}$  as an approximation to  $\mathbf{X}^{(t)}$ , where  $l$  is the sketch size. Take FD as an example, and we can formulate this sketching operation as

$$\mathbf{S}^{(t)} = \sqrt{\left( \mathbf{\Sigma}_l^{(t-1)} \right)^2 - \left( \sigma_l^{(t-1)} \right)^2 \mathbf{I}_l} \cdot \left( \mathbf{V}_l^{(t-1)} \right)^\top, \quad \mathbf{M}^{(t)} = \left( \mathbf{S}^{(t)} \left( \mathbf{S}^{(t)} \right)^\top + \lambda \mathbf{I}_l \right)^{-1}, \quad (3)$$

where  $\mathbf{\Sigma}_l^{(t-1)}$ ,  $\mathbf{V}_l^{(t-1)}$  are the result of rank- $l$  SVD on round  $t-1$  and  $\mathbf{M}^{(t)}$  is a diagonal matrix which can be stored efficiently. According to Woodbury's identity, we can rewrite the inverse of the sketched covariance matrix as

$$\left( \hat{\mathbf{A}}^{(t)} \right)^{-1} = \left( \lambda \mathbf{I}_d + \left( \mathbf{S}^{(t)} \right)^\top \mathbf{S}^{(t)} \right)^{-1} = \frac{1}{\lambda} \left( \mathbf{I}_d - \left( \mathbf{S}^{(t)} \right)^\top \mathbf{M}^{(t)} \mathbf{S}^{(t)} \right). \quad (4)$$

The computation in equation 2 requires  $\Omega(d^2)$  time. To improve efficiency, sketch-based methods replace  $\left( \mathbf{A}^{(t)} \right)^{-1}$  with  $\left( \hat{\mathbf{A}}^{(t)} \right)^{-1}$ . Notably,  $\left( \hat{\mathbf{A}}^{(t)} \right)^{-1}$  can be updated implicitly using the sketch matrix  $\mathbf{S}^{(t)}$  and  $\mathbf{M}^{(t)}$  in equation 4. Since matrix-vector multiplications with  $\mathbf{S}^{(t)}$  take  $O(ld)$  time and matrix-matrix multiplications with  $\mathbf{M}^{(t)}$  take  $O(l^2)$  time, the computation involving the inverse of the sketched covariance matrix is accelerated from  $\Omega(d^2)$  to  $O(ld + l^2)$ .

162 2.2 THE MOTIVATION OF REVISITING  
163

164 Without loss of generality, our analysis below is based on the original FD. To better illustrate our  
165 motivation, we first present the complete regret bound for the linear bandit using FD (Kuzborskij  
166 et al., 2019).

167 Let  $\bar{\sigma} = \sum_{t=1}^T (\sigma_l^{(t)})^2$  denote the sum of singular values reduced by FD up to the  $T$ -th round, where  
168  $l$  is the sketch size and  $(\sigma_l^{(t)})^2$  is the shrinking value at round  $t$ . According to Liberty (2013), it can  
169 be bounded by the spectral error  $\Delta_T$ , i.e.,  $\bar{\sigma} \leq \Delta_T$ , where  
170

$$\Delta_T := \min_{k < l} \frac{\|\mathbf{X}^{(T)} - \mathbf{X}_{[k]}^{(T)}\|_F^2}{l - k}. \tag{5}$$

171 Consequently, the regret of the sketch-based linear bandit can be formulated as  
172

$$\text{Regret}_T = \tilde{O} \left( (1 + \Delta_T)^{\frac{3}{2}} (l + d \log(1 + \Delta_T)) \sqrt{T} \right). \tag{6}$$

173 Denote  $k_*$  as the minimizer of equation 5. Ignoring logarithmic terms, we assume  $\Delta_T = T^\gamma$ . When  
174  $\gamma > \frac{1}{3}$ , the regret will exceed  $O(T)$ . More precisely, when the spectral tail  $\|\mathbf{X}^{(T)} - \mathbf{X}_{[k_*]}^{(T)}\|_F^2 =$   
175  $\Omega \left( (l - k_*)^{\frac{2}{3}} T^{\frac{1}{3}} \right)$ , the invalid linear regret will emerge.  
176

177 The above analysis highlights a key pitfall of sketch-based linear bandits: the pre-set sketch size is  
178 crucial. Next, we will explain, from both theoretical and experimental perspectives, why this pitfall is  
179 widespread and difficult to avoid in current methods. We present the following theorem to show that  
180 the spectral tail of the covariance matrix does not decrease rapidly, even in a non-adversarial setting.

181 **Theorem 1.** Suppose the chosen arm  $\mathbf{x}_t \in \mathbb{R}^d$  at round  $t$  is a random vector drawn iid from any  
182 distribution over  $r \leq d$  orthonormal vectors  $\mathbf{A}$ . For any sketch size  $l \leq r$ , the bound on the expected  
183 regret of linear bandits using FD is  $\Omega(T^2)$ .  
184

185 The detailed proof is provided in Appendix A. The-  
186 orem 1 shows that when the sketch size is insuffi-  
187 cient to capture most of the spectral information,  
188 sketch-based linear bandit methods will suffer from  
189 linear regret. Furthermore, as illustrated in Figure 1,  
190 we observe that an incorrect selection of the pre-set  
191 sketch size can significantly degrade performance.

192 Current methods use the single-scale sketching  
193 technique to approximate the covariance matrix.  
194 However, since the covariance matrix is determined  
195 in an online manner, its spectral information is un-  
196 known beforehand. Even if poor performance is  
197 detected during learning, adjusting the sketch size  
198 is not possible because the shrinking process in  
199 single-scale sketching is irreversible. Consequently,  
200 the pitfall of linear regret is difficult to avoid in  
201 previous works.  
202

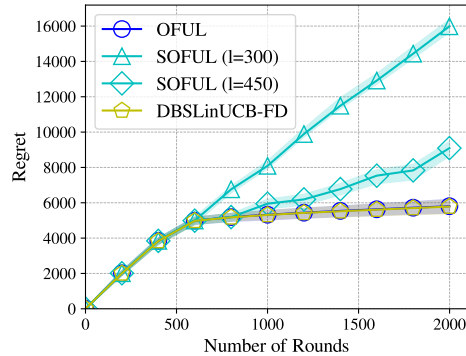


Figure 1: Regret of SOFUL, OFUL, and our method on synthetic data (details in section 5.2) The regret of SOFUL is nearly linear when sketch size  $l = 300$ .

210 3 DYADIC BLOCK SKETCHING FOR CONSTRAINED GLOBAL ERROR BOUND  
211

212 In this section, we propose a multi-scale sketching method called Dyadic Block Sketching, which  
213 provides a constrained global error bound for matrix sketching. This scalable method converts any  
214 streaming sketch into a matrix sketch with a constrained global error bound. It can adaptively adjust  
215 the sketch size in a streaming environment, tracking the best rank- $k$  approximation of the target  
matrix at a minimal cost. We employ the Frequent Directions sketch to illustrate this method.

### 3.1 DECOMPOSABILITY

Dyadic Block Sketching leverages the decomposability of streaming matrix sketches. Intuitively, decomposability means that if the rows of a matrix are divided into submatrices, sketches can be constructed for each submatrix, and these sketches can then be combined to form a sketch of the original matrix. In this process, the error of approximating the original matrix is the sum of the errors of approximating each submatrix. The following lemma demonstrates that all matrix sketches offering a covariance error guarantee exhibit decomposability.

**Lemma 1** (Decomposability). *Given a matrix  $\mathbf{X} \in \mathbb{R}^{n \times d}$ , we decompose  $\mathbf{X}$  into  $p$  submatrices, i.e.,  $\mathbf{X}^\top = [\mathbf{X}_1^\top, \mathbf{X}_2^\top, \dots, \mathbf{X}_p^\top]$ . For any  $i \in [p]$ , if we construct a matrix sketch with covariance error guarantee  $\epsilon_i$  for each sub-matrix  $\mathbf{X}_i$ , denoted as  $\mathbf{S}_i$ , such that  $\|\mathbf{X}_i^\top \mathbf{X}_i - \mathbf{S}_i^\top \mathbf{S}_i\|_2 \leq \epsilon_i \cdot \|\mathbf{X}_i - \mathbf{X}_{i[k]}\|_F^2$ . Then  $\mathbf{S}^\top = [\mathbf{S}_1^\top, \mathbf{S}_2^\top, \dots, \mathbf{S}_p^\top]$  is an approximation of  $\mathbf{X}^\top$  and the error bound is*

$$\|\mathbf{X}^\top \mathbf{X} - \mathbf{S}^\top \mathbf{S}\|_2 \leq \sum_{i=1}^p \epsilon_i \cdot \|\mathbf{X}_i - \mathbf{X}_{i[k]}\|_F^2.$$

*Proof.* Since we have  $\mathbf{X}^\top \mathbf{X} = \sum_{i=1}^p \mathbf{X}_i^\top \mathbf{X}_i$  and  $\mathbf{S}^\top \mathbf{S} = \sum_{i=1}^p \mathbf{S}_i^\top \mathbf{S}_i$ . Therefore

$$\|\mathbf{X}^\top \mathbf{X} - \mathbf{S}^\top \mathbf{S}\|_2 \leq \sum_{i=1}^p \|\mathbf{X}_i^\top \mathbf{X}_i - \mathbf{S}_i^\top \mathbf{S}_i\|_2 \leq \sum_{i=1}^p \epsilon_i \cdot \|\mathbf{X}_i - \mathbf{X}_{i[k]}\|_F^2,$$

and the Lemma follows.  $\square$

### 3.2 ALGORITHM DESCRIPTIONS

**High-Level Ideas.** The high-level idea is illustrated in Figure 2. We establish a logarithmic number of sketch sizes, each partitioning the stream into blocks. The sketch size of each subsequent block is double that of the previous one, thereby halving the maximum error caused by sketching. By maintaining a streaming sketch for each block, we can concatenate all the sketches to approximate the streaming matrix, with the error bounded by the decomposability property.

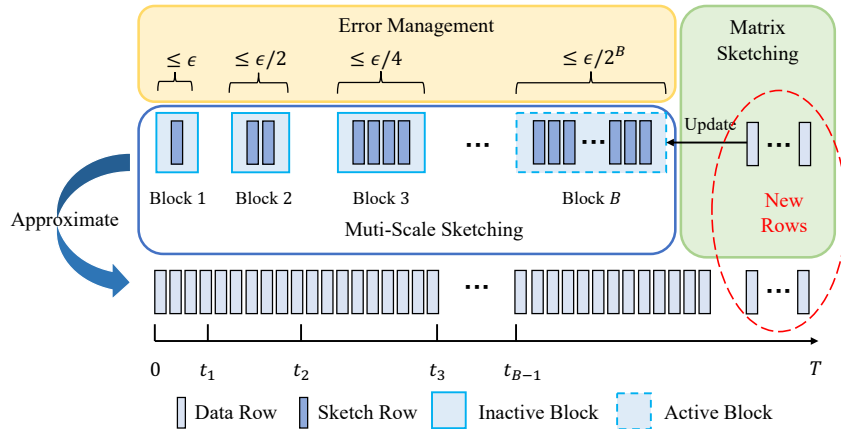


Figure 2: An illustration for Dyadic Block Sketching. For inactive Block  $i \in [B - 1]$ , the matrix sketch covers the data from  $t_{i-1}$  to  $t_i$ . For the active Block  $B$ , matrix sketching updates are performed on the new rows. We then merge the multi-scale matrix sketches to approximate the entire stream.

**Algorithm.** We start by defining the data structure for our algorithm. The matrix rows in the stream are divided into blocks, each covering a segment of consecutive, non-overlapping rows. The list of blocks is denoted as  $\mathcal{B}$ . For  $i = 1, 2, \dots$ , each block  $\mathcal{B}[i]$  is associated with a streaming sketch of size  $l$  (length) and block size (size). The size of block  $\mathcal{B}[i]$  is defined as the sum of the squared norms of the rows covered by  $\mathcal{B}[i]$ , specifically,  $\mathcal{B}[i].size = \sum_{\mathbf{x} \in \mathcal{B}[i]} \|\mathbf{x}\|_2^2$ .

We further define two states for the blocks: active and inactive. An active block receives updates, while an inactive block remains entirely fixed. As illustrated in Figure 2, there is always exactly one active block in the stream. Additionally, three key invariants must be maintained:

1. For each inactive block, the length must be greater than the rank of the rows it contains.
2. The sum of sketch rows stored in blocks should be less than  $d$ .
3. The size of block  $\mathcal{B}[i]$  should be less than  $l_0\epsilon$ , where  $l_0$  represents the initial sketch size and  $\epsilon$  is the error parameter.

Algorithm 1 presents the pseudo-code of Dyadic Block Sketching. When a new row  $\mathbf{x}_t$  is received, we first verify the maintenance of Invariant 2 (see Line 5). If the block count reaches its upper limit, the error from the streaming sketch becomes intolerable, necessitating the full preservation of the streaming rows' information. Therefore, we execute a complete rank-1 update on the sketch matrix.

---

**Algorithm 1: Dyadic Block Sketching**


---

**Input:** Data stream  $\{\mathbf{x}_t\}_{t=1}^T$ , sketch size  $l_0$ , error parameter  $\epsilon$ , regularization parameters  $\lambda$ , method  $\text{Sk} \in \{\text{FD}, \text{RFD}\}$

**Output:** Sketch matrix  $\mathbf{S}^{(t)}$ ,  $\mathbf{M}^{(t)}$

- 1 Initialize  $\mathcal{B}[0].size = 0$ ,  $\mathcal{B}[0].length = l_0$ ,  $B = 0$
- 2 Initialize  $\mathcal{B}[0].sketch$  by method  $\text{Sk}$
- 3 **for**  $t \leftarrow 1, \dots, T$  **do**
- 4     Receive  $\mathbf{x}_t$
- 5     **if**  $B \geq \lfloor \log(d/l_0 + 1) \rfloor - 1$  **then**
- 6         Update  $(\mathbf{S}^{(t)})^\top \mathbf{S}^{(t)} = (\mathbf{S}^{(t-1)})^\top \mathbf{S}^{(t-1)} + \mathbf{x}_t^\top \mathbf{x}_t$  and  $\mathbf{M}^{(t)}$  using rank-1 modifications
- 7     **else**
- 8         **if**  $\mathcal{B}[B].size + \|\mathbf{x}_t\|^2 > \epsilon \cdot \mathcal{B}[0].length$  and  $\mathcal{B}[B].length < rank$  **then**
- 9             Initialize  $\mathcal{B}[B+1].size = 0$ ,  $\mathcal{B}[B+1].length = 2 \times \mathcal{B}[B].length$
- 10            Initialize  $\mathcal{B}[B+1].sketch$  by method  $\text{Sk}$
- 11            Set  $B = B + 1$
- 12         Update  $\mathcal{B}[B].sketch$  with  $\mathbf{x}_t$
- 13         Set  $\mathbf{S}_B, \mathbf{M}_B, rank \leftarrow \mathcal{B}[B].sketch$
- 14         Update  $\mathcal{B}[B].size += \|\mathbf{x}_t\|^2$
- 15         Initialize empty matrix  $\mathbf{S}^{(t)}, \mathbf{M}^{(t)}$
- 16         **for**  $i \leftarrow 0, \dots, B$  **do**
- 17             Set  $\mathbf{S}_i, \mathbf{M}_i \leftarrow \mathcal{B}[i].sketch$
- 18             Combine  $\mathbf{S}^{(t)}$  with  $\mathbf{S}_i$ ,  $\mathbf{M}^{(t)}$  with  $\mathbf{M}_i$  as equation 7

---

In Lines 8 – 11, we control the errors to ensure the maintenance of Invariant 3. If the size of the active block exceeds the specified limit, we store the current block's information and create a new block with double the previous length to prevent further errors.

In Lines 12 – 14, we update the active block's information with  $\mathbf{x}_t$ . During this process, we can query the sketch matrices  $\mathbf{S}_B$  and  $\mathbf{M}_B$  in the active block (details in Appendix D). Additionally, the shrinkage of the deterministic sketch provides us with the block's current rank if the sketch size exceeds the block's rank. We use the variable  $rank$  to track this value. If the shrinking value is non-zero, we set  $rank$  to  $\mathcal{B}[B].length$ ; otherwise, we assign  $rank$  to the block's rank.

In Lines 15 – 18, we query the sketch of the entire stream. To retrieve the sketch matrix  $\mathbf{S}^{(t)}$  and  $\mathbf{M}^{(t)}$ , we combine them with the previous matrices as follows

$$\mathbf{S}^{(t)} = \begin{pmatrix} \mathbf{S}^{(t)} \\ \mathbf{S}_i \end{pmatrix}, \quad \mathbf{M}^{(t)} = \left( \begin{pmatrix} \mathbf{M}^{(t)} & \mathbf{S}^{(t)}(\mathbf{S}_i)^\top \\ \mathbf{S}_i(\mathbf{S}^{(t)})^\top & \mathbf{M}_i \end{pmatrix} + \lambda \mathbf{I} \right)^{-1}. \quad (7)$$

Since the inactive blocks remain fixed, we can store the combined result of the sketch matrix in the inactive blocks. In practice, we can avoid the looped calculation of equation 7 and perform the combination only with the active block.

### 3.3 ANALYSIS

We explore the error guarantees along with the space and time complexities of the Dyadic Block Sketching method. Initially, we prove a general theorem that establishes the relationship between the complexities of the Dyadic Block Sketching algorithm and the streaming sketch utilized in each block. Subsequently, we examine how a specific deterministic streaming sketch can be integrated into the Dyadic Block Sketching framework. We present a theorem detailing the space usage and update costs associated with Dyadic Block Sketching. The detailed proof is provided in Appendix B.

**Theorem 2.** *Suppose a streaming matrix sketch, denoted as  $\kappa$ , achieves a covariance error  $\|\mathbf{X}^\top \mathbf{X} - \mathbf{S}^\top \mathbf{S}\|_2 \leq \eta \cdot \|\mathbf{X}\|_F^2$  with  $\ell_\eta$  rows and  $\mu_\eta$  update time. Applying  $\kappa$  as the sketching method for each block in the Dyadic Block Sketching and  $l_0$  is the initial sketch size, we generate a matrix sketch  $\mathbf{S}$  for the entire streaming matrix  $\mathbf{X}$  with an error guarantee  $\|\mathbf{X}^\top \mathbf{X} - \mathbf{S}^\top \mathbf{S}\|_2 \leq 2\epsilon$ . Assuming  $\text{rank}(\mathbf{X}) = k$  and that the rows are normalized, the space cost for Dyadic Block Sketching is  $O\left(d \cdot \sum_{i=0}^B \ell_{\frac{1}{2^i l_0}}\right)$ , and the update cost is  $O\left(\mu_{\frac{1}{2^B l_0}}\right)$ , where  $B = \left\lceil \min\left\{\log \frac{k}{l_0}, \frac{T}{\epsilon l_0}\right\} \right\rceil + 1$ .*

Note that different streaming sketches will result in varying costs. To illustrate this, we provide a corollary using the well-known deterministic sketching method Frequent Directions (FD).

**Dyadic Block Sketching for FD.** This algorithm employs the FD sketch for each block in the Dyadic Block Sketching framework. Recall that with a given error parameter  $\eta$ , the FD sketch requires a space of  $\ell_\eta = O(1/\eta)$  and processes updates at an amortized cost of  $\mu_\eta = O(d/\eta)$ . As outlined in Theorem 2, we derive the following corollary:

**Corollary 1.** *The Dyadic Block Sketching algorithm for FD uses  $O(2dk - dl_0)$  space and processes an update with  $O(dk)$  amortized cost.*

**Remark 1** (Efficient Implementation). *The primary computational costs of the algorithm include calculating the SVD to obtain  $\mathbf{S}^{(t)}$  and performing matrix multiplication to compute  $\mathbf{M}^{(t)}$ , both of which cost  $O(dl^2)$ , where  $l$  is the current sketch size. However, the amortized update cost can be effectively reduced from  $O(dl^2)$  to  $O(dl)$  either by doubling the space, as detailed in Algorithm 3 in Appendix C, or by employing the Gu-Eisenstat procedure (Gu & Eisenstat, 1993).*

**Remark 2** (Worst-Case Analysis). *Compared to the single streaming sketch algorithm, our method effectively controls the global error of matrix approximation by limiting error sizes within each block, thus enabling dynamic adjustment of the sketch size. Particularly when dealing with a full-rank matrix with a heavy spectral tail, sketching methods should be avoided to prevent the pitfall of linear regret. However, this control is impossible with a single sketch due to the irreversible nature of the shrinking process in sketching. In the worst-case scenario, our method ensures that the streaming matrix problem can be dynamically adjusted to revert to a non-sketch situation.*

## 4 LINEAR BANDITS THROUGH DYADIC BLOCK SKETCHING

In this section, we introduce a novel framework for efficient sketch-based linear bandits, termed DBSLinUCB, which leverages Dyadic Block Sketching. As outlined in Section 2.2, a key limitation of previous methods stems from their reliance on single-scale matrix sketching, resulting in a *space-bounded* linear bandit approach. The use of a fixed sketch size leads to uncontrollable spectral loss,  $\Delta_T$ , ultimately causing linear regret.

In contrast, Dyadic Block Sketching employs a multi-scale matrix sketching strategy, where the sketch size is adaptively adjusted based on the pre-set parameter  $\epsilon$ . Consequently, DBSLinUCB is an *error-bounded* linear bandit method that effectively addresses and overcomes the pitfalls of linear regret present in existing approaches.

The procedure, detailed in Algorithm 2, builds on prior sketch-based algorithms but incorporates the Dyadic Block Sketching method to effectively manage the error in approximating the covariance

**Algorithm 2: DBSLinUCB**


---

**Input:** Data stream  $\{\mathbf{x}_t\}_{t=1}^T$ , sketch size  $l_0$ , error parameter  $\epsilon$ , regularization parameters  $\lambda$ , method  $\text{Sk} \in \{\text{FD}, \text{RFD}\}$ , confidence  $\delta$

- 1 Initialize a Dyadic Block Sketching instance  $\text{sketch}(\mathbf{S}^{(0)}, \mathbf{M}^{(0)})$  with  $l_0, \lambda$ , method  $\text{Sk}$
- 2 **for**  $t \leftarrow 1, \dots, T$  **do**
- 3     Get arm set  $\mathcal{X}_t$
- 4     Compute the confidence ellipsoid  $\hat{\beta}_{t-1}(\delta)$  using equation 9
- 5     Select  $\mathbf{x}_t = \arg \max_{\mathbf{x} \in \mathcal{X}_t} \left\{ \mathbf{x}^\top \hat{\boldsymbol{\theta}}_{t-1} + \hat{\beta}_{t-1}(\delta) \cdot \|\mathbf{x}\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}} \right\}$
- 6     Receive the reward  $r_t$
- 7     Update  $\text{sketch}(\mathbf{S}^{(t)}, \mathbf{M}^{(t)})$  with  $\mathbf{x}_t$  and compute  $\hat{\mathbf{A}}^{(t)}$  and  $\hat{\boldsymbol{\theta}}_t$  using equation 8

---

matrix. At each round  $t$ , we employ the sketch matrix  $\mathbf{S}^{(t)}$  to approximate the covariance matrix, from which we derive the sketched regularized least squares estimator as follows

$$\hat{\mathbf{A}}^{(t)} = \mathbf{S}^{(t)\top} \mathbf{S}^{(t)} + \lambda \mathbf{I} \quad , \quad \hat{\boldsymbol{\theta}}_t = \left( \hat{\mathbf{A}}^{(t)} \right)^{-1} \sum_{s=1}^t r_s \mathbf{x}_s. \quad (8)$$

Denote  $\mathcal{X}_t$  as the set of arms available at round  $t$  and  $\hat{\beta}_{t-1}(\delta)$  as the confidence ellipsoid,  $l_{B_t}$  as the sketch size in the active block. The arms selected by DBSLinUCB are determined by solving the following constrained optimization problem:

$$\mathbf{x}_t = \arg \max_{\mathbf{x} \in \mathcal{X}_t} \max_{\boldsymbol{\theta} \in \mathbb{R}^d} \mathbf{x}^\top \boldsymbol{\theta} \quad \text{such that} \quad \left\| \boldsymbol{\theta} - \hat{\boldsymbol{\theta}} \right\|_{\hat{\mathbf{A}}^{(t-1)}} \leq \hat{\beta}_{t-1}(\delta). \quad (9)$$

The updates to the sketched regularized least squares estimator and the calculations for the confidence ellipsoid can be efficiently completed in  $O(d l_{B_t} + l_{B_t}^2)$  time using the Woodbury identity as stated in equation 4. This makes DBSLinUCB significantly more efficient than traditional linear bandit algorithms, which require  $\Omega(d^2)$  in both time and space.

DBSLinUCB represents a scalable framework for efficient sketch-based linear bandits that are capable of incorporating various streaming sketching techniques. We now explore two deterministic sketching techniques that provide different regret bounds of linear bandits.

**DBSLinUCB using FD.** We explore the Frequent Directions (FD) (see Algorithm 4), a deterministic sketching method (Liberty, 2013; Ghashami et al., 2016). FD uniquely maintains the invariant that the last row of the sketch matrix,  $\mathbf{S}$ , is always zero. In each round, a new row  $\mathbf{a}_t$  is inserted into this last row of  $\mathbf{S}$ , and the matrix undergoes singular value decomposition into  $\mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^\top$ . Subsequently,  $\mathbf{S}$  is updated to  $\sqrt{\boldsymbol{\Sigma}_l^2 - \sigma \mathbf{I}} \cdot \mathbf{V}_l^\top$ , where  $\sigma$  represents the square of the  $l$ -th singular value. Given that the rows of  $\mathbf{S}$  are orthogonal,  $\mathbf{M} = (\mathbf{S} \mathbf{S}^\top + \lambda \mathbf{I})^{-1}$  remains a diagonal matrix, facilitating efficient maintenance. We integrate FD into DBSLinUCB and established the following regret bound:

**Theorem 3.** *Assume that  $\|\boldsymbol{\theta}_*\|_2 \leq H$ ,  $\|\mathbf{x}\|_2 \leq L$ , and  $L \geq \sqrt{\lambda}$ . Suppose that the noise is conditionally  $R$ -subgaussian, where  $R$  is a fixed constant. The sketch size in the active block at round  $t$  is denoted as  $l_{B_t}$ . Given the error parameter  $\epsilon$ , then with a probability of  $1 - \frac{1}{T}$ , the regret of Algorithm 2 utilizing  $\text{Sk} = \text{FD}$  is*

$$\text{Regret}_T \stackrel{\bar{O}}{=} \frac{L(R + H\sqrt{\lambda})}{\sqrt{\lambda}} \cdot \left( d \ln \left( 1 + \frac{\epsilon}{\lambda} \right) + 2l_{B_T} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} \right) \right) \cdot \left( 1 + \frac{\epsilon}{\lambda} \right)^{\frac{3}{2}} \sqrt{T}.$$

**Remark 3.** *Compared to linear bandits that use a single FD sketch, our approach relies on a predetermined error parameter  $\epsilon$  rather than fixed sketch size  $l$ . This ensures a sublinear regret bound of order  $\tilde{O}(\sqrt{T})$  without requiring prior knowledge of the streaming matrix. The sketch size in DBSLinUCB can be dynamically adjusted to match the desired order of regret. In practice, if the goal is to significantly enhance efficiency at the expense of a higher order of regret,  $\epsilon$  can be set as a function of  $T$ . Another advantage of DBSLinUCB is that the initial sketch size is independent of the error term  $\epsilon$ ; it only affects the total computational and space cost.*



**DBSLinUCB using RFD.** We employ the Robust Frequent Directions (RFD) sketch (see Algorithm 5), a sketching strategy designed to address the rank deficiency issue inherent in FD (Luo et al., 2019). RFD reduces the approximation error of FD by maintaining a counter  $\alpha$ , which quantifies the spectral error. More precisely, RFD employs  $\mathbf{S}^\top \mathbf{S} + \alpha \mathbf{I}$  to approximate  $\mathbf{A}^\top \mathbf{A}$ . We integrate RFD into DBSLinUCB and established the following regret bound:

**Theorem 4.** Assume that  $\|\boldsymbol{\theta}_*\|_2 \leq H$ ,  $\|\mathbf{x}\|_2 \leq L$ , and  $L \geq \sqrt{\lambda}$ . Suppose that the noise is conditionally  $R$ -subgaussian, where  $R$  is a fixed constant. The sketch size in the active block at round  $t$  is denoted as  $l_{B_t}$ . Given the error parameter  $\epsilon$ , then with a probability of  $1 - \frac{1}{T}$ , the regret of Algorithm 2 utilizing  $\text{Sk} = \text{RFD}$  is

$$\text{Regret}_T \stackrel{\text{O}}{=} \frac{L}{\sqrt{\lambda}} \cdot \sqrt{\ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} + \left( 1 - \frac{2 - B_T}{2^{B_T+1}} \right) \cdot \frac{\epsilon}{\lambda} \right)} \cdot \left( H \cdot \sqrt{\lambda} + \epsilon + R \cdot \sqrt{d \ln \left( 1 + \frac{\epsilon}{\lambda} \right) + 2l_{B_T} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} + \left( 1 - \frac{2 - B_T}{2^{B_T+1}} \right) \cdot \frac{\epsilon}{\lambda} \right)} \right).$$

**Remark 4.** Compared to DBSLinUCB using FD, the order of the error term  $\epsilon$  is reduced from  $3/2$  to  $1/2$ . Apart from logarithmic terms, decoupling the dimensions  $d$  and  $\epsilon$  further reduces the impact of the error. These properties indicate that Dyadic Block Sketching maintains the excellent characteristics of RFD. In fact, despite having the same error bound as DBSLinUCB using FD, DBSLinUCB using RFD satisfies both positive definite monotonicity and well-conditioned properties. We present details and proof in Appendix G.

## 5 EXPERIMENTS

In this section, we empirically verify the efficiency and effectiveness of our algorithms. We conduct experiments on both synthetic and real-world datasets. Each experiment is performed over 20 different random permutations of the datasets. All experiments are performed on a machine with 24-core Intel(R) Xeon(R) Gold 6240R 2.40GHz CPU and 256 GB memory.

### 5.1 MATRIX APPROXIMATION

We evaluate the performance of the proposed Dyadic Block Sketching in terms of matrix approximation. We compare it with FD (Liberty, 2013). We generated a synthetic dataset with  $n = 1250$  rows and  $d = 100$  columns. Specifically, each row  $\mathbf{a}_t \in \mathbb{R}^{100}$  is independently drawn from a multivariate Gaussian distribution  $\mathbf{a}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)$ . We set the sketch size  $l_0 = 50$  for FD. We set the initial sketch size  $l_0 = 16$  and the error parameter  $\epsilon = 2000$  for Dyadic Block Sketching.

Figure 3 shows the spectral norm error  $\|\mathbf{A}_t^\top \mathbf{A}_t - \mathbf{S}_t^\top \mathbf{S}_t\|_2$  and its upper bound for matrix sketching, where  $\mathbf{A}_t$  is the steaming matrix at round  $t$  and  $\mathbf{S}_t$  is the skech matrix at round  $t$ . We observe that Dyadic Block Sketching provides a constrained global error bound for matrix sketching. Compared with FD, the rate of error growth in Dyadic Block Sketching decreases over time, effectively limiting the linear growth of the spectral tail.

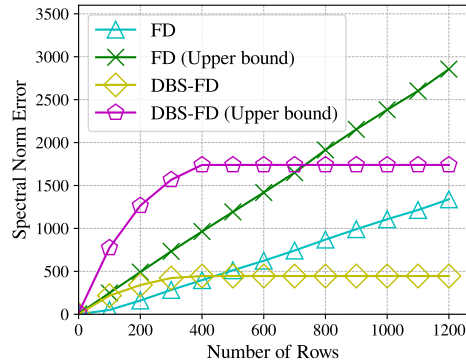


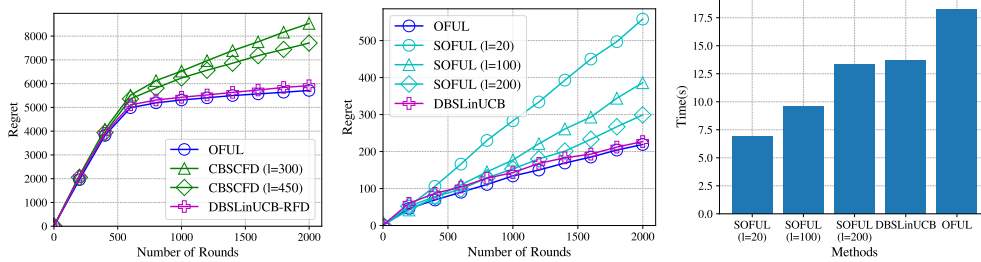
Figure 3: Comparison among FD and our DBS-FD w.r.t. the error and its upper bound

### 5.2 ONLINE REGRESSION IN SYNTHETIC DATA

In this section, we evaluate our DBSLinUCB on synthetic datasets. The baselines include the non-sketched method OFUL (Abbasi-Yadkori et al., 2011) and the sketch-based methods SOFUL (Kuzborskij et al., 2019), CBSCFD (Chen et al., 2021). Inspired by the experimental settings in Chen et al. (2021), we build synthetic datasets using multivariate Gaussian distributions  $\mathcal{N}(0, \mathbf{I}_d)$

with 100 arms and  $d = 500$  features per context. The true parameter  $\theta_*$  is drawn from  $\mathcal{N}(0, \mathbf{I}_d)$  and is normalized. The confidence ellipsoid  $\beta$  of all algorithms is searched in  $\{10^{-4}, 10^{-3}, \dots, 1\}$  and  $\lambda$  is searched in  $\{2 \times 10^{-4}, 2 \times 10^{-3}, \dots, 2 \times 10^4\}$ . We set the sketch size  $l = 300, 450$  for SOFUL and CBSCFD and the initial sketch size  $l_0 = 64$  for DBSLinUCB. Additionally, we set the error parameter  $\epsilon = 2000$  for DBSLinUCB.

Experimental results in Figure 1 (in Section 2.2), 4a show that DBSLinUCB using FD and RFD consistently outperforms the other sketch-based algorithms in terms of the regret of online learning. We observe that when  $l = 300$ , SOFUL and CBSCFD perform significantly worse than DBSLinUCB, with SOFUL exhibiting nearly linear regret. Moreover, DBSLinUCB achieves sublinear regret similar to OFUL by providing a constrained global error bound. Our experimental results confirm our analysis in Section 2.2, indicating that for all existing sketch-based linear bandit algorithms, inappropriate sketch size selection can lead to the pitfall of linear regret.



(a) Regret on synthetic data (b) Regret on MNIST (c) Running time on MNIST

Figure 4: (a): Cumulative regret of the compared algorithms, the proposed DBSLinUCB using RFD on a synthetic dataset; (b), (c): Cumulative regret and total running time of the compared algorithms, the proposed DBSLinUCB using FD on MNIST

### 5.3 ONLINE CLASSIFICATION IN REAL-WORLD DATA

We perform online classification on the real-world dataset MNIST to evaluate the performance of our methods. The dataset contains 60,000 samples, each with  $d = 784$  features, and there are  $M = 10$  possible labels for each sample. We follow the experimental setup in Kuzborskij et al. (2019). Specifically, we construct the online classification problem within the contextual bandit setting as follows: given a dataset with data in  $M$  labels, we first choose one cluster as the target label. In each round, we randomly draw one sample from each label and compose an arm set of  $M$  samples in  $M$  contexts. The algorithms choose one sample from the arm set and observe the reward based on whether the selected sample belongs to the target label. The reward is 1 if the selected sample comes from the target label and 0 otherwise. We set sketch size  $l = 20, 100, 200$  for SOFUL and  $l_0 = 2$  for DBSLinUCB. We set the error parameter  $\epsilon = 1000$  for DBSLinUCB. The choice of confidence ellipsoid and regularization parameter follows the previous section.

Figures 4b and 4c compare the online mistakes and running times of different algorithms. Our findings indicate that, for a given dataset, there exists an optimal sketch size (e.g.,  $l = 200$ ) that captures most of the spectral information of the original matrix, thereby accelerating the algorithm without significantly compromising performance. However, selecting this optimal sketch size for SOFUL is challenging due to the lack of prior knowledge about the data. When  $l = 20$  or  $l = 100$ , the regret of SOFUL is significantly worse than that of the non-sketched method, OFUL. In contrast, DBSLinUCB matches the performance of OFUL by adaptively adjusting the sketch size to the near-optimal value of  $l = 200$  while being significantly faster than OFUL.

## 6 CONCLUSION

This paper addresses the current pitfall of linear regret in sketch-based linear bandits for the first time. We propose Dyadic Block Sketching with a constrained global error bound and provide formal theoretical guarantees. By leveraging Dyadic Block Sketching, we present a framework for efficient sketch-based linear bandits. Even in the worst-case scenario, our method can achieve sublinear regret without prior knowledge of the covariance matrix. Extensive experimental evaluations on real and synthetic datasets demonstrate the excellent performance and efficiency of our methods.

## REFERENCES

- 540  
541  
542 Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic  
543 bandits. *Advances in neural information processing systems*, 24, 2011.
- 544  
545 Dimitris Achlioptas. Database-friendly random projections. In *Proceedings of the twentieth ACM*  
546 *SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pp. 274–281, 2001.
- 547  
548 Peter Auer. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine*  
*Learning Research*, 3(Nov):397–422, 2002.
- 549  
550 Daniele Calandriello, Alessandro Lazaric, and Michal Valko. Efficient second-order online kernel  
551 learning with adaptive embedding. *Advances in Neural Information Processing Systems*, 30, 2017.
- 552  
553 Daniele Calandriello, Luigi Carratino, Alessandro Lazaric, Michal Valko, and Lorenzo Rosasco.  
554 Gaussian process optimization with adaptive sketching: Scalable and no regret. In *Conference on*  
*Learning Theory*, pp. 533–557. PMLR, 2019.
- 555  
556 Cheng Chen, Luo Luo, Weinan Zhang, Yong Yu, and Yijiang Lian. Efficient and robust high-  
557 dimensional linear contextual bandits. In *Proceedings of the Twenty-Ninth International Conference*  
*on International Joint Conferences on Artificial Intelligence*, pp. 4259–4265, 2021.
- 558  
559 Wei Chu, Lihong Li, Lev Reyzin, and Robert Schapire. Contextual bandits with linear payoff  
560 functions. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and*  
561 *Statistics*, pp. 208–214. JMLR Workshop and Conference Proceedings, 2011.
- 562  
563 Varsha Dani, Sham M Kakade, and Thomas Hayes. The price of bandit information for online  
564 optimization. *Advances in Neural Information Processing Systems*, 20, 2007.
- 565  
566 Amit Deshpande and Luis Rademacher. Efficient volume sampling for row/column subset selection.  
567 In *2010 IEEE 51st annual symposium on foundations of computer science*, pp. 329–338. IEEE,  
2010.
- 568  
569 Yash Deshpande and Andrea Montanari. Linear bandits in high dimension and recommendation  
570 systems. In *2012 50th Annual Allerton Conference on Communication, Control, and Computing*  
*(Allerton)*, pp. 1750–1754. IEEE, 2012.
- 571  
572 Alan Frieze, Ravi Kannan, and Santosh Vempala. Fast monte-carlo algorithms for finding low-rank  
573 approximations. *Journal of the ACM (JACM)*, 51(6):1025–1041, 2004.
- 574  
575 Mina Ghashami, Edo Liberty, Jeff M Phillips, and David P Woodruff. Frequent directions: Simple  
576 and deterministic matrix sketching. *SIAM Journal on Computing*, 45(5):1762–1792, 2016.
- 577  
578 Alon Gonen, Francesco Orabona, and Shai Shalev-Shwartz. Solving ridge regression using sketched  
579 preconditioned svrg. In *International conference on machine learning*, pp. 1397–1405. PMLR,  
2016.
- 580  
581 Ming Gu and Stanley C Eisenstat. A stable and fast algorithm for updating the singular value  
582 decomposition, 1993.
- 583  
584 Ilja Kuzborskij, Leonardo Cella, and Nicolò Cesa-Bianchi. Efficient linear bandits through matrix  
585 sketching. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pp.  
586 177–185. PMLR, 2019.
- 587  
588 Edo Liberty. Simple and deterministic matrix sketching. In *Proceedings of the 19th ACM SIGKDD*  
*international conference on Knowledge discovery and data mining*, pp. 581–588, 2013.
- 589  
590 Haipeng Luo, Alekh Agarwal, Nicolo Cesa-Bianchi, and John Langford. Efficient second order  
591 online learning by sketching. *Advances in Neural Information Processing Systems*, 29, 2016.
- 592  
593 Luo Luo, Cheng Chen, Zhihua Zhang, Wu-Jun Li, and Tong Zhang. Robust frequent directions with  
application in online learning. *Journal of Machine Learning Research*, 20(45):1–41, 2019.
- Jayadev Misra and David Gries. Finding repeated elements. *Science of computer programming*, 2(2):  
143–152, 1982.

594 Herbert Robbins. Some aspects of the sequential design of experiments. 1952.  
595  
596 Tamas Sarlos. Improved approximation algorithms for large matrices via random projections. In  
597 *2006 47th annual IEEE symposium on foundations of computer science (FOCS'06)*, pp. 143–152.  
598 IEEE, 2006.

599 David P Woodruff et al. Sketching as a tool for numerical linear algebra. *Foundations and Trends®*  
600 *in Theoretical Computer Science*, 10(1–2):1–157, 2014.  
601

602 Xiao Xu, Fang Dong, Yanghua Li, Shaojian He, and Xin Li. Contextual-bandit based personalized  
603 recommendation with time-varying user interests. In *Proceedings of the AAAI Conference on*  
604 *Artificial Intelligence*, volume 34, pp. 6518–6525, 2020.

605 Xiaotian Yu, Michael R Lyu, and Irwin King. Cbrap: Contextual bandits with random projection. In  
606 *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.  
607

608 Xiao Zhang, Ninglu Shao, Zihua Si, Jun Xu, Wenhan Wang, Hanjing Su, and Ji-Rong Wen. Reward  
609 imputation with sketching for contextual batched bandits. *Advances in Neural Information*  
610 *Processing Systems*, 36, 2024.  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647

## A PROOF OF THEOREM 1

For linear bandits using FD, at round  $t$ , the sketch matrix  $\mathbf{S}^{(t-1)} \in \mathbb{R}^{l \times d}$  is utilized, and a new row  $\mathbf{x}_t$  is received, where  $l$  denotes the sketch size. Let  $\Sigma = \mathbb{E}[\mathbf{x}_t \mathbf{x}_t^\top]$  be the covariance matrix of the arms, and let  $\lambda_i$  be the  $i$ -th eigenvalue of  $\Sigma$ . By definition,  $\mathbf{x}_t$  follows the distribution  $\mathbf{x}_t = \mathbf{a}_i$  with probability  $\lambda_i$ , where  $\mathbf{a}_i$  is the  $i$ -th vector of  $\mathbf{A}$ .

We first consider the expected value of the shrinking factor when the sketch is full-rank. If  $\text{rank}(\mathbf{S}^{(t-1)}) = l - 1$ . Let  $\mathbf{S}^{(t-1)} = \mathbf{U}\Sigma\mathbf{V}$  be the SVD of  $\mathbf{S}^{(t-1)}$ , and  $\mathbf{v}_i$  be the  $i$ -th row of  $\mathbf{V}$ . Denote the set of basis vectors not in the row space of  $\mathbf{S}^{(t-1)}$  as  $W_{t-1} = \mathbf{A} \setminus \{\mathbf{v}_1, \dots, \mathbf{v}_{l-1}\}$ , we have  $|W_{t-1}| = r - l + 1$ . If  $\mathbf{x}_t \in \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_{l-1})$ , we have the shrinking value  $(\sigma_l^{(t)})^2 = 0$ ; otherwise  $(\sigma_l^{(t)})^2 = 1$ , with probability  $\sum_{\mathbf{a}_i \in W_{t-1}} \lambda_i \geq \sum_{i=l}^r \lambda_i$ . Therefore, If  $\text{rank}(\mathbf{S}^{(t-1)}) = l - 1$ , we have

$$\mathbb{E} \left[ (\sigma_l^{(t)})^2 \mid \mathbf{S}^{(t-1)} \right] \geq \sum_{i=l}^r \lambda_i.$$

If  $\text{rank}(\mathbf{S}^{(t-1)}) < l - 1$ , this means that  $\mathbf{S}^{(t-1)}$  contains fewer than  $l - 1$  distinct vectors drawn from  $\mathbf{A}$ . Let  $I_i$  be the indicator variable for drawing  $\mathbf{a}_i$  in the first  $t - 1$  rounds. Then we have the expected number of distinct vectors at round  $t - 1$

$$\mathbb{E} \left[ \sum_{i=1}^r I_i \right] = \sum_{i=1}^r (1 - (1 - \lambda_i)^{t-1}).$$

Using Markov's inequality, we have

$$\text{Prob} \left[ r - \sum_{i=1}^r I_i \geq r - l + 2 \right] \leq \frac{\sum_{i=1}^r 1 - (1 - \lambda_i)^{t-1}}{r - l + 2} \leq \frac{r(1 - \lambda_r)^{t-1}}{2}.$$

Note that this is precisely the probability of having fewer than  $l - 1$  distinct vectors in the first  $t - 1$  rounds. We conclude that for  $t \geq \log \left( \frac{r}{\lambda_r} \right) + 1$ ,  $\text{Prob}[\text{rank}(\mathbf{S}_{t-1}) = l - 1] \geq \frac{1}{2}$ . This implies that  $\mathbb{E}[\rho_t] \geq \frac{1}{2} \sum_{i=l}^r \lambda_i$  after an initial logarithmic number of rounds.

Therefore, assuming  $T \geq 2 \log \left( \frac{r}{\lambda_r} \right)$ , the expected accumulated shrinking value is at least:

$$\mathbb{E} \left[ \sum_{t=1}^T (\sigma_l^{(t)})^2 \right] \geq \frac{T}{4} \sum_{i=l}^r \lambda_i.$$

Note that the accumulated shrinking value is upper-bounded by the spectral loss. According to the regret in equation 6, we conclude that the regret upper bound is  $\Omega(T^2)$  in expectation.

## B PROOF OF THEOREM 2

We begin our proof by considering the number of blocks. Let the entire data stream be denoted as  $\mathbf{X}^\top = [\mathbf{X}_0^\top, \mathbf{X}_1^\top, \dots, \mathbf{X}_B^\top]$ . Block  $i$  covers the submatrix  $\mathbf{X}_i$  and stores the corresponding matrix sketch  $\mathbf{S}_i$ .

According to Invariant 3, each submatrix  $\mathbf{X}_i$  in block  $i$  contains at least  $\lfloor \epsilon l_0 \rfloor$  rows. Since there are  $T$  rows available for allocation in the entire matrix, the maximum number of blocks is  $\left\lceil \frac{T}{\epsilon l_0} \right\rceil$ .

When  $\epsilon$  is small, the block length grows exponentially. By Invariant 1, the length of the last active block will be exactly greater than  $k$ , i.e.,  $2^{B-1}l_0 \leq k \leq 2^B l_0$ .

Thus, we take the minimum of the two scenarios to determine the number of blocks, resulting in  $B = \left\lceil \min \left\{ \log \frac{k}{l_0}, \frac{T}{\epsilon l_0} \right\} \right\rceil + 1$ .

For the error guarantee, recall that the error comprises two components: the error from the active block and the error from merging  $B$  inactive blocks. Given that the streaming sketch can detect the best rank- $k$  approximation as long as the sketch size exceeds  $k$ , the approximation error in the  $B$ -th active block is effectively zero.

By maintaining Invariant 3, the size of the  $i$ -th inactive block is bounded by  $l_0 \epsilon$ . Additionally, the  $i$ -th block employs a streaming matrix sketch with an error parameter of  $\frac{1}{2^i l_0}$ , thereby ensuring that the maximum error introduced by a sketch at the  $i$ -th block is at most  $\frac{\epsilon}{2^i}$ .

Combining all  $B + 1$  sketches, we use  $\mathbf{S}^\top = [\mathbf{S}_0^\top, \mathbf{S}_1^\top, \dots, \mathbf{S}_B^\top]$  to approximate  $\mathbf{X}$ . By Lemma 1, this provides the following error guarantee for the entire streaming matrix:

$$\begin{aligned} \|\mathbf{X}^\top \mathbf{X} - \mathbf{S}^\top \mathbf{S}\|_2 &\leq \sum_{i=0}^B \|\mathbf{X}_i^\top \mathbf{X}_i - \mathbf{S}_i^\top \mathbf{S}_i\|_2 \\ &= \sum_{i=0}^{B-1} \|\mathbf{X}_i^\top \mathbf{X}_i - \mathbf{S}_i^\top \mathbf{S}_i\|_2 + \|\mathbf{X}_B^\top \mathbf{X}_B - \mathbf{S}_B^\top \mathbf{S}_B\|_2 \\ &\leq \sum_{i=0}^{B-1} \frac{\epsilon}{2^i} + 0 \\ &\leq 2\epsilon. \end{aligned}$$

For space usage, the  $i$ -th block employs a streaming matrix sketch with an error parameter of  $\frac{1}{2^i l_0}$ , resulting in a sketch of size  $\ell_{\frac{1}{2^i l_0}}$ . Therefore, the total number of sketched rows is  $\sum_{i=0}^B \ell_{\frac{1}{2^i l_0}}$ .

Consequently, the total space requirement is  $O\left(d \cdot \sum_{i=0}^B \ell_{\frac{1}{2^i l_0}}\right)$ .

For the update cost, since only the active sketch requires updating, the cost is  $O\left(\mu_{\frac{1}{2^B l_0}}\right)$ .

## C FAST ALGORITHM OF DYADIC BLOCK SKETCHING

The computational cost of FD and RFD, as detailed in Algorithm 4 and 5, is primarily driven by the singular value decomposition (SVD) operations. At round  $t$ , with  $B_t + 1$  blocks, we denote  $l_i$  as the sketch size for the  $i$ -th block. It incurs an amortized time of  $O(dl_{B_t}^2)$  due to standard SVD processes in the active block. Additionally, the operation to compute  $\mathbf{M}^{(t)}$  via matrix multiplication and matrix inversion also requires  $O\left(\sum_{i=0}^{B_t-1} l_i \cdot l_{B_t} \cdot d + \left(\sum_{i=0}^{B_t} l_i\right)^3\right) = O(dl_{B_t}^2)$ . We can enhance the efficiency of our Dyadic Block Sketching by doubling the sketch size, as detailed in Algorithm 3.

Notice that within each epoch, the update of  $\mathbf{M}^{(t)}$  can be formulate as

$$\mathbf{M}^{(t)} = \begin{pmatrix} \mathbf{M}^{(t-1)} + \frac{\phi \phi^\top}{\xi} & \frac{-\phi}{\xi} \\ \frac{-\phi^\top}{\xi} & \frac{1}{\xi} \end{pmatrix}, \quad (10)$$

where  $\phi = \mathbf{M}^{(t-1)} \mathbf{S}^{(t-1)} \mathbf{x}_t^\top$  and  $\xi = \mathbf{x}_t \mathbf{x}_t^\top - \mathbf{x}_t (\mathbf{S}^{(t-1)})^\top \phi + \alpha + \lambda$ .

When the method Sk = FD,  $\alpha$  is set to 0; conversely, when the method Sk = RFD,  $\alpha$  serves as the counter maintained in the RFD sketch.

Given that the length of  $\mathbf{M}^{(t)}$  is at most twice the length of  $\mathcal{B}[B].length$ , the amortized computation time required for  $\mathbf{M}^{(t)}$  is limited to  $O(dl_{B_t})$ . Additionally, we perform the SVD only after every addition of  $\mathcal{B}[B].length$  rows, reducing the amortized update time complexity to  $O(dl_{B_t})$ .

**Algorithm 3: Fast Dyadic Block Sketching**


---

**Input:** Data stream  $\{\mathbf{x}_t\}_{t=1}^T$ , sketch size  $l_0$ , error parameter  $\epsilon$ , regularization parameters  $\lambda$ , method  $\text{Sk} \in \text{FD, RFD}$

**Output:** Sketch matrix  $\mathbf{S}^{(t)}, \mathbf{M}^{(t)}$

- 1 Initialize  $\mathcal{B}[0].size = 0, \mathcal{B}[0].length = l_0, B = 0$
- 2 Initialize  $\mathcal{B}[0].sketch$  by method  $\text{Sk}$
- 3 **for**  $t \leftarrow 1, \dots, T$  **do**
- 4     Receive  $\mathbf{x}_t$
- 5     **if**  $B \geq \lceil \log(d/l_0 + 1) \rceil - 1$  **then**
- 6         Update  $(\mathbf{S}^{(t)})^\top \mathbf{S}^{(t)} = (\mathbf{S}^{(t-1)})^\top \mathbf{S}^{(t-1)} + \mathbf{x}_t^\top \mathbf{x}_t$
- 7         Update  $\mathbf{M}^{(t)}$  using rank-1 modifications
- 8     **else**
- 9         **if**  $\mathcal{B}[B].size + \|\mathbf{x}_t\|^2 > \frac{\epsilon}{2} \cdot \mathcal{B}[0].length$  and  $\mathcal{B}[B].length < rank$  **then**
- 10             Initialize  $\mathcal{B}[B+1].size = 0, \mathcal{B}[B+1].length = 2 \times \mathcal{B}[B].length$
- 11             Initialize  $\mathcal{B}[B+1].sketch$  by method  $\text{Sk}$
- 12             Set  $B \leftarrow B + 1$
- 13         Append  $\mathbf{x}_t$  below  $\mathcal{B}[B].sketch$
- 14         Update  $\mathcal{B}[B].size += \|\mathbf{x}_t\|^2$
- 15         Initialize empty matrix  $\mathbf{S}, \mathbf{M}$
- 16         **for**  $i \leftarrow 0, \dots, B - 1$  **do**
- 17             Set  $\mathbf{S}_i, \mathbf{M}_i \leftarrow \mathcal{B}[i].sketch$
- 18             Combine  $\mathbf{S}$  with  $\mathbf{S}_i$
- 19             Combine  $\mathbf{M}$  with  $\mathbf{M}_i$
- 20         **if**  $\mathcal{B}[B].sketch$  have  $2 \cdot \mathcal{B}[B].length$  rows **then**
- 21             Update  $\mathcal{B}[B].sketch$  by method  $\text{Sk}$
- 22             Set  $\mathbf{S}_B, \mathbf{M}_B, rank \leftarrow \mathcal{B}[B].sketch$
- 23             Update  $\mathbf{S}^{(t)} = \begin{pmatrix} \mathbf{S} \\ \mathbf{S}_B \end{pmatrix}, \mathbf{M}^{(t)} = \left( \begin{pmatrix} \mathbf{M} & \mathbf{S}(\mathbf{S}_B)^\top \\ \mathbf{S}_B \mathbf{S}^\top & \mathbf{M}_B \end{pmatrix} + \lambda \mathbf{I} \right)^{-1}$
- 24         **else**
- 25             Update  $\mathbf{S}^{(t)} = \begin{pmatrix} \mathbf{S} \\ \mathbf{S}_B \end{pmatrix}$
- 26             Update  $\mathbf{M}^{(t)}$  by equation 10

---

**D PSEUDO-CODE OF DETERMINISTIC MATRIX SKETCHING****Algorithm 4: FD sketch**


---

**Input:** Data  $\mathbf{X} \in \mathbb{R}^{T \times d}$ , sketch size  $l$ , regularization  $\lambda$

**Output:** Sketch  $\mathbf{S}, \mathbf{M}$

- 1 Initialize  $\mathbf{S} \leftarrow \mathbf{0}^{l \times d}, \mathbf{M} \leftarrow \frac{1}{\lambda} \mathbf{I}_l$
- 2 **for**  $t \leftarrow 1, \dots, T$  **do**
- 3     Append  $\mathbf{x}_t$  to the last row of  $\mathbf{S}$
- 4     Compute  $[\mathbf{U}, \boldsymbol{\Sigma}, \mathbf{V}] \leftarrow \text{svd}(\mathbf{S})$
- 5     Set  $\sigma \leftarrow \sigma_l^2$
- 6     Update  $\mathbf{S} \leftarrow \sqrt{\boldsymbol{\Sigma}_l^2 - \sigma \mathbf{I}} \cdot \mathbf{V}_l^\top$
- 7     Update
- 8          $\mathbf{M} \leftarrow \text{diag} \left\{ \frac{1}{\lambda + \sigma_1^2 - \sigma}, \dots, \frac{1}{\lambda} \right\}$

---

**Algorithm 5: RFD sketch**


---

**Input:** Data  $\mathbf{X} \in \mathbb{R}^{T \times d}$ , sketch size  $l$ , regularization  $\lambda$

**Output:** Sketch  $\mathbf{S}, \mathbf{M}$  and counter  $\alpha$

- 1 Initialize  $\mathbf{S} \leftarrow \mathbf{0}^{l \times d}, \mathbf{M} \leftarrow \frac{1}{\lambda} \mathbf{I}_l, \alpha \leftarrow 0$
- 2 **for**  $t \leftarrow 1, \dots, T$  **do**
- 3     Append  $\mathbf{x}_t$  to the last row of  $\mathbf{S}$
- 4     Compute  $[\mathbf{U}, \boldsymbol{\Sigma}, \mathbf{V}] \leftarrow \text{svd}(\mathbf{S})$
- 5     Set  $\sigma \leftarrow \sigma_l^2, \alpha \leftarrow \alpha + \sigma$
- 6     Update  $\mathbf{S} \leftarrow \sqrt{\boldsymbol{\Sigma}_l^2 - \sigma \mathbf{I}} \cdot \mathbf{V}_l^\top$
- 7     Set
- 8          $\mathbf{M} \leftarrow \text{diag} \left\{ \frac{1}{\lambda + \sigma_1^2 - \sigma + \alpha}, \dots, \frac{1}{\lambda + \alpha} \right\}$

---

The pseudo-code for deterministic matrix sketching methods is displayed in Algorithm 4 and Algorithm 5. Note that these deterministic matrix sketching methods can be accelerated by doubling the sketch size. More details can be found in Appendix C.

## E PROOF OF THEOREM 3

Following Abbasi-Yadkori et al. (2011), we structure the proof in two parts: first, we establish bounds for the approximate confidence ellipsoid, and second, we delineate the bounds for regret. Denote  $B_t$  as the number of blocks at round  $t$ , and  $\bar{\sigma}_i$  as the sum of shrinking singular values in the sketch of block  $i$ . Let  $l_{B_t}$  be the sketch size in the active block at round  $t$ . We start by establishing an intermediate result concerning the confidence ellipsoid.

**Theorem 5.** *Let  $\hat{\theta}_t$  be the RLS estimate constructed by an arbitrary policy for linear bandits after  $t$  rounds of play. For any  $\delta \in (0, 1)$ , the optimal unknown weight  $\theta_*$  belongs to the set  $\Theta_t \equiv \left\{ \theta \in \mathbb{R}^d : \left\| \theta - \hat{\theta}_t \right\|_{\hat{A}^{(t)}} \leq \hat{\beta}_t(\delta) \right\}$  with probability at least  $1 - \delta$ , where*

$$\begin{aligned} \hat{\beta}_t(\delta) = & R \cdot \sqrt{1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda}} \cdot \sqrt{2 \ln \left( \frac{1}{\delta} \right) + d \ln \left( 1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_t} \cdot \ln \left( 1 + \frac{tL^2}{2l_{B_t}\lambda} \right)} \\ & + H \cdot \frac{\lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i}{\sqrt{\lambda}}. \end{aligned}$$

*Proof.* According to Algorithm 2, the approximate covariance matrix is

$$\hat{A}^{(t)} = \lambda \mathbf{I} + \sum_{i=1}^{B_t} \left( \mathbf{S}_i^{(t)} \right)^\top \mathbf{S}_i^{(t)},$$

where  $\mathbf{S}_i^{(t)}$  is the sketch matrix in block  $i$  at round  $t$ . Define  $\mathbf{z}_1^\top, \dots, \mathbf{z}_t^\top \in \mathbb{R}^d$  is the noise sequence conditionally R-subgaussian for a fixed constant  $R$  and  $\mathbf{r}_t^\top = (r_1, r_2, \dots, r_t) \in \mathbb{R}^d$  is the reward vector. We begin by noticing that

$$\hat{\theta}_t = \left( \hat{A}^{(t)} \right)^{-1} \mathbf{X}_t^\top \mathbf{r}_t = \left( \hat{A}^{(t)} \right)^{-1} \mathbf{X}_t^\top (\mathbf{X}_t \theta_* + \mathbf{z}_t).$$

Therefore, we can decompose  $\left\| \hat{\theta}_t - \theta_* \right\|_{\hat{A}^{(t)}}^2$  into two parts as follows

$$\begin{aligned} & \left\| \hat{\theta}_t - \theta_* \right\|_{\hat{A}^{(t)}}^2 \\ &= \left( \hat{\theta}_t - \theta_* \right)^\top \hat{A}^{(t)} \left( \hat{\theta}_t - \theta_* \right) \\ &= \left( \hat{\theta}_t - \theta_* \right)^\top \hat{A}^{(t)} \left( \left( \hat{A}^{(t)} \right)^{-1} \mathbf{X}_t^\top (\mathbf{X}_t \theta_* + \mathbf{z}_t) - \theta_* \right) \\ &= \underbrace{\left( \hat{\theta}_t - \theta_* \right)^\top \hat{A}^{(t)} \left( \left( \hat{A}^{(t)} \right)^{-1} \mathbf{X}_t^\top \mathbf{X}_t \theta_* - \theta_* \right)}_{\text{Term 1: Bias Error}} + \underbrace{\left( \hat{\theta}_t - \theta_* \right)^\top \mathbf{X}_t^\top \mathbf{z}_t}_{\text{Term 2: Variance Error}}. \end{aligned} \tag{11}$$

**Bounding the bias error.** We first focus on bounding the first term. We have that

$$\begin{aligned} & \left( \hat{\theta}_t - \theta_* \right)^\top \hat{A}^{(t)} \left( \left( \hat{A}^{(t)} \right)^{-1} \mathbf{X}_t^\top \mathbf{X}_t \theta_* - \theta_* \right) \\ &= \left( \hat{\theta}_t - \theta_* \right)^\top \left( \hat{A}^{(t)} \right)^{\frac{1}{2}} \left( \hat{A}^{(t)} \right)^{-\frac{1}{2}} \left( \mathbf{X}_t^\top \mathbf{X}_t \theta_* - \hat{A}^{(t)} \theta_* \right) \\ &= \left( \hat{\theta}_t - \theta_* \right)^\top \left( \hat{A}^{(t)} \right)^{\frac{1}{2}} \left( \hat{A}^{(t)} \right)^{-\frac{1}{2}} \left[ \left( \mathbf{A}^{(t)} - \hat{A}^{(t)} \right) \theta_* - \lambda \theta_* \right]. \end{aligned} \tag{12}$$



864 In accordance with the decomposability of matrix sketches, as detailed in Lemma 1, we have

$$865 \left\| \mathbf{X}_t^\top \mathbf{X}_t - \sum_{i=1}^{B_t} \left( \mathbf{S}_i^{(t)} \right)^\top \mathbf{S}_i^{(t)} \right\|_2 \leq \sum_{i=1}^{B_t} \bar{\sigma}_i \quad (13)$$

866 By Cauchy-Schwartz inequality and the triangle inequality, we have

$$867 \begin{aligned} & \left( \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right)^\top \left( \hat{\mathbf{A}}^{(t)} \right)^{\frac{1}{2}} \left( \hat{\mathbf{A}}^{(t)} \right)^{-\frac{1}{2}} \left[ \left( \mathbf{A}^{(t)} - \hat{\mathbf{A}}^{(t)} \right) \boldsymbol{\theta}_* - \lambda \boldsymbol{\theta}_* \right] \\ & \leq \left| \lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i \right| \cdot \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}} \cdot \left\| \boldsymbol{\theta}_* \right\|_{\left( \hat{\mathbf{A}}^{(t)} \right)^{-1}} \\ & \leq H \cdot \frac{\lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i}{\sqrt{\lambda}} \cdot \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}, \end{aligned} \quad (14)$$

868 where the last inequality holds because  $\hat{\mathbf{A}}^{(t)} \succeq \lambda \mathbf{I}$  and  $\left\| \boldsymbol{\theta}_* \right\|_2 \leq H$ .

869 **Bounding the variance error.** Then, we aim to bound the second term. We use the following self-normalized martingale concentration inequality by Abbasi-Yadkori et al. (2011).

870 **Proposition 1.** *Assume that  $\mathbf{z}_1, \dots, \mathbf{z}_t$  is a conditionally  $R$ -subgaussian real-valued stochastic process and  $\mathbf{X}_t^\top = [\mathbf{x}_1^\top, \dots, \mathbf{x}_t^\top]$  is any stochastic process such that  $\mathbf{x}_i$  is measurable with respect to the  $\sigma$ -algebra generated by  $\mathbf{z}_1, \dots, \mathbf{z}_t$ . Then, for any  $\delta > 0$ , with probability at least  $1 - \delta$ , for all  $t \geq 0$ ,*

$$871 \left\| \mathbf{X}_t^\top \mathbf{z}_t \right\|_{\left( \mathbf{A}^{(t)} \right)^{-1}}^2 \leq 2R^2 \ln \left( \frac{1}{\delta} \left| \mathbf{A}^{(t)} \right|^{\frac{1}{2}} \left| \lambda \mathbf{I} \right|^{-\frac{1}{2}} \right).$$

872 Notice that the variance error can be reformulated as

$$873 \begin{aligned} \left( \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right)^\top \mathbf{X}_t^\top \mathbf{z}_t &= \left( \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right)^\top \left( \mathbf{A}^{(t)} \right)^{-\frac{1}{2}} \left( \mathbf{A}^{(t)} \right)^{\frac{1}{2}} \mathbf{X}_t^\top \mathbf{z}_t \\ &\leq \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}} \cdot \frac{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\mathbf{A}^{(t)}}}{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}} \cdot \left\| \mathbf{X}_t^\top \mathbf{z}_t \right\|_{\left( \mathbf{A}^{(t)} \right)^{-1}}, \end{aligned} \quad (15)$$

874 where the last inequality uses Cauchy-Schwartz inequality.

875 For any vector  $\mathbf{a}$ , we have

$$876 \begin{aligned} \left\| \mathbf{a} \right\|_{\mathbf{A}^{(t)}}^2 - \left\| \mathbf{a} \right\|_{\hat{\mathbf{A}}^{(t)}}^2 &= \mathbf{a}^\top \left( \mathbf{A}^{(t)} - \hat{\mathbf{A}}^{(t)} \right) \mathbf{a} \\ &= \mathbf{a}^\top \left( \mathbf{X}^\top \mathbf{X} - \sum_{i=1}^{B_t} \left( \mathbf{S}_i^{(t)} \right)^\top \mathbf{S}_i^{(t)} \right) \mathbf{a} \\ &\leq \sum_{i=1}^{B_t} \bar{\sigma}_i \cdot \left\| \mathbf{a} \right\|_2^2. \end{aligned} \quad (16)$$

877 Therefore, the ratios of norms on the right-hand side of equation 15 can be bounded as

$$878 \begin{aligned} \frac{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\mathbf{A}^{(t)}}}{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}} &= \sqrt{\frac{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\mathbf{A}^{(t)}}^2}{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}^2}} \\ &\leq \sqrt{\frac{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}^2 + \sum_{i=1}^{B_t} \bar{\sigma}_i \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_2^2}{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}^2}} \\ &\leq \sqrt{1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda}}. \end{aligned} \quad (17)$$

Substituting equation 17 and Proposition 1 into equation 15 gives

$$\begin{aligned}
& \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}} \cdot \frac{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\mathbf{A}^{(t)}}}{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}} \cdot \left\| \mathbf{X}_t^\top \mathbf{z}_t \right\|_{(\mathbf{A}^{(t)})^{-1}} \\
& \leq \sqrt{1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda}} \cdot \sqrt{2R^2 \ln \left( \frac{1}{\delta} |\mathbf{A}^{(t)}|^{\frac{1}{2}} |\lambda \mathbf{I}|^{-\frac{1}{2}} \right)} \cdot \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}.
\end{aligned} \tag{18}$$

Motivated by Abbasi-Yadkori et al. (2011); Kuzborskij et al. (2019), we apply the multi-scale sketch-based determinant-trace inequality. Compared to the non-sketched version, this inequality depends on the approximate covariance matrix  $\hat{\mathbf{A}}$ , reflecting the costs associated with the shrinkage due to multi-scale sketching.

**Lemma 2.** For any  $t \geq 1$ , define  $\mathbf{A}^{(t)} = \lambda \mathbf{I} + \mathbf{X}_t^\top \mathbf{X}_t$ , and assume  $\|\mathbf{x}_t\|_2 \leq L$ , we have

$$\ln \left( \frac{|\mathbf{A}^{(t)}|}{|\lambda \mathbf{I}|} \right) \leq d \ln \left( 1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_t} \cdot \ln \left( 1 + \frac{tL^2}{2l_{B_t}\lambda} \right).$$

*Proof.*  $\sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)}$  has rank at most  $2l_{B_t}$  due to the Dyadic Block Sketching. Since  $\hat{\mathbf{A}}^{(t)} = \lambda \mathbf{I} + \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)}$  and  $\mathbf{A}^{(t)} \preceq \hat{\mathbf{A}}^{(t)} + \sum_{i=1}^{B_t} \bar{\sigma}_i \cdot \mathbf{I}$ , we have

$$\begin{aligned}
|\mathbf{A}^{(t)}| & \leq \left| \hat{\mathbf{A}}^{(t)} + \sum_{i=1}^{B_t} \bar{\sigma}_i \cdot \mathbf{I} \right| \\
& \leq \left( \lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i \right)^{d-2l_{B_t}} \cdot \left( \frac{\sum_{i=1}^{2l_{B_t}} \left( \lambda_i \left( \hat{\mathbf{A}}^{(t)} \right) + \sum_{i=1}^{B_t} \bar{\sigma}_i \right)}{2l_{B_t}} \right)^{2l_{B_t}} \\
& \leq \left( \lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i \right)^{d-2l_{B_t}} \cdot \left( \lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i + \frac{\text{Tr} \left( \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)} \right)}{2l_{B_t}} \right)^{2l_{B_t}} \\
& \leq \left( \lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i \right)^{d-2l_{B_t}} \cdot \left( \lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i + \frac{tL^2}{2l_{B_t}} \right)^{2l_{B_t}},
\end{aligned}$$

where the last inequality holds because

$$\begin{aligned}
& \text{Tr} \left( \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)} \right) \\
& \leq \text{Tr} \left( \mathbf{X}_t^\top \mathbf{X}_t \right) \\
& = \sum_{s=1}^t \mathbf{x}_s^\top \mathbf{x}_s \\
& \leq tL^2
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\ln \left( \frac{|\mathbf{A}^{(t)}|}{|\lambda \mathbf{I}|} \right) &\leq \ln \left\{ \left( \frac{\lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda} \right)^{d-2l_{B_t}} \cdot \left( \frac{\lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i + \frac{tL^2}{2l_{B_t}}}{\lambda} \right)^{2l_{B_t}} \right\} \\
&= (d - 2l_{B_t}) \ln \left( 1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_t} \ln \left( 1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda} + \frac{tL^2}{2l_{B_t}\lambda} \right) \\
&\leq d \ln \left( 1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_t} \cdot \ln \left( 1 + \frac{tL^2}{2l_{B_t}\lambda} \right).
\end{aligned}$$

□

According to Lemma 2, we finally bound the variance error term as follows

$$\begin{aligned}
&\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}} \cdot \frac{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\mathbf{A}^{(t)}}}{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}} \cdot \left\| \mathbf{X}_t^\top \mathbf{z}_t \right\|_{(\mathbf{A}^{(t)})^{-1}} \\
&\leq \sqrt{1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda}} \cdot \sqrt{2R^2 \ln \left( \frac{1}{\delta} |\mathbf{A}^{(t)}|^{\frac{1}{2}} |\lambda \mathbf{I}|^{-\frac{1}{2}} \right)} \cdot \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}} \\
&\leq R \cdot \sqrt{1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda}} \cdot \sqrt{2 \ln \left( \frac{1}{\delta} \right) + d \ln \left( 1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_t} \cdot \ln \left( 1 + \frac{tL^2}{2l_{B_t}\lambda} \right)} \cdot \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}.
\end{aligned}$$

Sum up the bias error and the variance error and divide both sides of equation 11 by  $\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}$  simultaneously, we have

$$\begin{aligned}
\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}} &\leq R \cdot \sqrt{1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda}} \cdot \sqrt{2 \ln \left( \frac{1}{\delta} \right) + d \ln \left( 1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_t} \cdot \ln \left( 1 + \frac{tL^2}{2l_{B_t}\lambda} \right)} \\
&\quad + H \cdot \frac{\lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i}{\sqrt{\lambda}},
\end{aligned}$$

which concludes the proof. □

Having established the confidence ellipsoid, we now focus on analyzing the regret. We begin with an analysis of the instantaneous regret.

Recall that the optimal arm at round  $t$  is defined as  $\mathbf{x}_t^* = \arg \max_{\mathbf{x} \in \mathcal{X}_t} (\mathbf{x}^\top \boldsymbol{\theta}_*)$ . On the other hand, the principle of optimism in the face of uncertainty ensures that  $(\mathbf{x}_t, \hat{\boldsymbol{\theta}}_{t-1}) = \arg \max_{(\mathbf{x}, \boldsymbol{\theta}) \in \mathcal{X}_t \times \Theta_{t-1}} \mathbf{x}^\top \boldsymbol{\theta}$ . By

denoting  $\tilde{\boldsymbol{\theta}}_t$  as the RLS estimator, we utilize these facts to establish the bound on the instantaneous regret as follows

$$\begin{aligned}
&(\mathbf{x}_t^* - \mathbf{x}_t)^\top \boldsymbol{\theta}_* \\
&\leq \mathbf{x}_t^\top \hat{\boldsymbol{\theta}}_{t-1} - \mathbf{x}_t^\top \boldsymbol{\theta}_* \\
&= \mathbf{x}_t^\top (\hat{\boldsymbol{\theta}}_{t-1} - \tilde{\boldsymbol{\theta}}_{t-1}) + \mathbf{x}_t^\top (\tilde{\boldsymbol{\theta}}_{t-1} - \boldsymbol{\theta}_*) \\
&\leq \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}} \cdot \left( \left\| \hat{\boldsymbol{\theta}}_{t-1} - \tilde{\boldsymbol{\theta}}_{t-1} \right\|_{\hat{\mathbf{A}}^{(t-1)}} + \left\| \tilde{\boldsymbol{\theta}}_{t-1} - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t-1)}} \right) \\
&\leq 2\hat{\beta}_{t-1}(\delta) \cdot \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}}.
\end{aligned} \tag{19}$$

Now, we are prepared to establish the upper bound of regret. Utilizing equation 19 and Cauchy-Schwartz inequality, we derive the following bound

$$\begin{aligned}
\text{Regret}_T &= \sum_{t=1}^T \max_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\top \boldsymbol{\theta}_* - \sum_{t=1}^T \mathbf{x}_t^\top \boldsymbol{\theta}_* \\
&\leq 2 \sum_{t=1}^T \min \left\{ HL, \hat{\beta}_{t-1}(\delta) \cdot \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}} \right\} \\
&\leq 2 \sum_{t=1}^T \hat{\beta}_{t-1}(\delta) \min \left\{ \frac{L}{\sqrt{\lambda}}, \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}} \right\} \\
&\leq 2 \cdot \max \left\{ 1, \frac{L}{\sqrt{\lambda}} \right\} \cdot \hat{\beta}_T(\delta) \cdot \sum_{t=1}^T \min \left\{ 1, \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}} \right\} \\
&\leq 2 \cdot \max \left\{ 1, \frac{L}{\sqrt{\lambda}} \right\} \cdot \hat{\beta}_T(\delta) \cdot \sqrt{T \sum_{t=1}^T \min \left\{ 1, \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}}^2 \right\}}.
\end{aligned} \tag{20}$$

We further bound the terms in the above. In particular, we formulate  $\hat{\beta}_T(\delta)$  by Theorem 5 as follows

$$\begin{aligned}
\hat{\beta}_T(\delta) &= R \sqrt{1 + \frac{\sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda}} \cdot \sqrt{2 \ln \frac{1}{\delta} + d \ln \left( 1 + \frac{\sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_T} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} \right)} \\
&\quad + H\sqrt{\lambda} \left( 1 + \frac{\sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda} \right).
\end{aligned} \tag{21}$$

Besides, we adopt the Sketched leverage scores established by Kuzborskij et al. (2019) as follows

**Proposition 2** (Lemma 6 of Kuzborskij et al. (2019)). *The sketched leverage scores through sketching at round  $T$  can be upper bounded as*

$$\begin{aligned}
&\sum_{t=1}^T \min \left\{ 1, \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t)})^{-1}}^2 \right\} \\
&\leq 2 \left( 1 + \frac{\sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda} \right) \cdot \ln \left( \frac{|\mathbf{A}^{(T)}|}{|\lambda \mathbf{I}|} \right) \\
&\leq 2 \left( 1 + \frac{\sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda} \right) \cdot \left( d \ln \left( \frac{1 + \sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_t} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} \right) \right).
\end{aligned}$$

Combining equation 21, equation 20 and Proposition 2, assume  $L \geq \sqrt{\lambda}$ , we have

$$\begin{aligned}
\text{Regret}_T &\leq 2 \cdot \max \left\{ 1, \frac{L}{\sqrt{\lambda}} \right\} \cdot \hat{\beta}_T(\delta) \cdot \sqrt{T \sum_{t=1}^T \min \left\{ 1, \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}}^2 \right\}} \\
&\stackrel{\tilde{O}}{=} \frac{L}{\sqrt{\lambda}} \cdot \sqrt{T} \cdot \left( 1 + \frac{\sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda} \right) \cdot \left( d \ln \left( 1 + \frac{\sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_T} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} \right) \right) \\
&\quad \cdot \left( R \sqrt{1 + \frac{\sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda}} \cdot \sqrt{2 \ln \frac{1}{\delta} + d \ln \left( 1 + \frac{\sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_T} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} \right)} \right. \\
&\quad \left. + H \sqrt{\lambda} \left( 1 + \frac{\sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda} \right) \right) \\
&\stackrel{\tilde{O}}{=} \frac{L(R + H\sqrt{\lambda})}{\sqrt{\lambda}} \cdot \left( d \ln \left( 1 + \frac{\sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_T} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} \right) \right) \\
&\quad \cdot \left( 1 + \frac{\sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda} \right)^{\frac{3}{2}} \sqrt{T}.
\end{aligned}$$

According to Theorem 2, we can bound the spectral error by error  $\epsilon$ , which is

$$\text{Regret}_T \stackrel{\tilde{O}}{=} \frac{L(R + H\sqrt{\lambda})}{\sqrt{\lambda}} \cdot \left( d \ln \left( 1 + \frac{\epsilon}{\lambda} \right) + 2l_{B_T} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} \right) \right) \cdot \left( 1 + \frac{\epsilon}{\lambda} \right)^{\frac{3}{2}} \sqrt{T}.$$

## F PROOF OF THEOREM 4

We denote  $B_t$  as the number of blocks at round  $t$ , and  $\bar{\sigma}_i$  as the cumulative shrinking singular values in the sketch of block  $i$ . Let  $l_{B_t}$  be the sketch size in the active block at round  $t$ . Similarly, our analysis establishes an intermediate result regarding the confidence ellipsoid.

**Theorem 6.** *Let  $\hat{\theta}_t$  be the RLS estimate constructed by an arbitrary policy for linear bandits after  $t$  rounds of play. For any  $\delta \in (0, 1)$ , the optimal unknown weight  $\theta_*$  belongs to the set  $\Theta_t \equiv \left\{ \theta \in \mathbb{R}^d : \left\| \theta - \hat{\theta}_t \right\|_{\hat{\mathbf{A}}^{(t)}} \leq \hat{\beta}_t(\delta) \right\}$  with probability at least  $1 - \delta$ , where*

$$\begin{aligned}
\hat{\beta}_t(\delta) &= R \cdot \sqrt{2 \ln \left( \frac{1}{\delta} \right) + d \ln \left( 1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_t} \cdot \ln \left( 1 + \frac{tL^2}{2l_{B_t}\lambda} + \frac{h_t}{\lambda} \right)} \\
&\quad + H \cdot \sqrt{\lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i}
\end{aligned}$$

and

$$h_t = \sum_{i=1}^{B_t} \bar{\sigma}_i - \frac{\sum_{i=1}^{B_t} l_i \cdot \bar{\sigma}_i}{2l_{B_t}}.$$

*Proof.* Notice that RFD uses the adaptive regularization term to approximate the covariance matrix, i.e.,  $\hat{\mathbf{A}}^{(t)} = \lambda \mathbf{I} + \sum_{i=1}^{B_t} \alpha_i^{(t)} \mathbf{I} + \sum_{i=1}^{B_t} \left( \mathbf{S}_i^{(t)} \right)^\top \mathbf{S}_i^{(t)}$ , where  $\mathbf{S}_i^{(t)}$  is the sketch matrix in block  $i$  and  $\alpha_i^{(t)}$  is the adaptive regularization term of RFD at round  $t$ .

1134 Similarly, we decompose  $\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}^2$  into two parts as follows

$$\begin{aligned}
1135 & \\
1136 & \\
1137 & \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}^2 \\
1138 & = \left( \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right)^\top \hat{\mathbf{A}}^{(t)} \left( \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right) \\
1139 & = \left( \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right)^\top \hat{\mathbf{A}}^{(t)} \left( \left( \hat{\mathbf{A}}^{(t)} \right)^{-1} \mathbf{X}_t^\top (\mathbf{X}_t \boldsymbol{\theta}_* + \mathbf{z}_t) - \boldsymbol{\theta}_* \right) \\
1140 & \\
1141 & = \underbrace{\left( \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right)^\top \hat{\mathbf{A}}^{(t)} \left( \left( \hat{\mathbf{A}}^{(t)} \right)^{-1} \mathbf{X}_t^\top \mathbf{X}_t \boldsymbol{\theta}_* - \boldsymbol{\theta}_* \right)}_{\text{Term 1: Bias Error}} + \underbrace{\left( \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right)^\top \mathbf{X}_t^\top \mathbf{z}_t}_{\text{Term 2: Variance Error}}. \\
1142 & \\
1143 & \\
1144 & \\
1145 & \\
1146 & 
\end{aligned}$$

1147 **Bounding the bias error.** For the bias error term, we have

$$\begin{aligned}
1148 & \left( \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right)^\top \hat{\mathbf{A}}^{(t)} \left( \left( \hat{\mathbf{A}}^{(t)} \right)^{-1} \mathbf{X}_t^\top \mathbf{X}_t \boldsymbol{\theta}_* - \boldsymbol{\theta}_* \right) \\
1149 & = \left( \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right)^\top \left( \hat{\mathbf{A}}^{(t)} \right)^{\frac{1}{2}} \left( \hat{\mathbf{A}}^{(t)} \right)^{-\frac{1}{2}} \left( \mathbf{X}_t^\top \mathbf{X}_t \boldsymbol{\theta}_* - \hat{\mathbf{A}}^{(t)} \boldsymbol{\theta}_* \right) \\
1150 & \\
1151 & = \left( \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right)^\top \left( \hat{\mathbf{A}}^{(t)} \right)^{\frac{1}{2}} \left( \hat{\mathbf{A}}^{(t)} \right)^{-\frac{1}{2}} \left( \mathbf{X}_t^\top \mathbf{X}_t - \lambda \mathbf{I} - \sum_{i=1}^{B_t} \alpha_i^{(t)} \mathbf{I} - \sum_{i=1}^{B_t} \left( \mathbf{S}_i^{(t)} \right)^\top \mathbf{S}_i^{(t)} \right) \boldsymbol{\theta}_* \\
1152 & \\
1153 & \triangleq \left( \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right)^\top \left( \hat{\mathbf{A}}^{(t)} \right)^{\frac{1}{2}} \left( \hat{\mathbf{A}}^{(t)} \right)^{-\frac{1}{2}} \mathbf{D}_t \cdot \boldsymbol{\theta}_* \\
1154 & \\
1155 & \\
1156 & \\
1157 & 
\end{aligned} \tag{22}$$

1158 Since  $\mathbf{D}_t = \mathbf{X}_t^\top \mathbf{X}_t - \lambda \mathbf{I} - \sum_{i=1}^{B_t} \alpha_i^{(t)} \mathbf{I} - \sum_{i=1}^{B_t} \left( \mathbf{S}_i^{(t)} \right)^\top \mathbf{S}_i^{(t)}$ , for any unit vector  $\mathbf{a}$ , we have

$$\begin{aligned}
1160 & \\
1161 & \left| \mathbf{a}^\top \mathbf{D}_t \mathbf{a} \right| = \left| \mathbf{a}^\top \left( \mathbf{X}_t^\top \mathbf{X}_t - \lambda \mathbf{I} - \sum_{i=1}^{B_t} \alpha_i^{(t)} \mathbf{I} - \sum_{i=1}^{B_t} \left( \mathbf{S}_i^{(t)} \right)^\top \mathbf{S}_i^{(t)} \right) \mathbf{a} \right| \\
1162 & \\
1163 & = \left| \mathbf{a}^\top \left( \mathbf{X}_t^\top \mathbf{X}_t - \sum_{i=1}^{B_t} \left( \mathbf{S}_i^{(t)} \right)^\top \mathbf{S}_i^{(t)} \right) \mathbf{a} - \lambda \mathbf{I} - \sum_{i=1}^{B_t} \alpha_i^{(t)} \mathbf{I} \right|. \\
1164 & \\
1165 & \\
1166 & \\
1167 & 
\end{aligned} \tag{23}$$

1168 According to Theorem 2, we can get

$$1169 \quad 0 \leq \mathbf{a}^\top \left( \mathbf{X}_t^\top \mathbf{X}_t - \sum_{i=1}^{B_t} \left( \mathbf{S}_i^{(t)} \right)^\top \mathbf{S}_i^{(t)} \right) \mathbf{a} \leq \sum_{i=1}^{B_t} \bar{\sigma}_i.$$

1171 Bring the above equation into equation 23, since  $\sum_{i=1}^{B_t} \alpha_i^{(t)} = \sum_{i=1}^{B_t} \bar{\sigma}_i$ , we can bound the spectral

1172 norm of  $\mathbf{D}_t$  as follows

$$1173 \quad \|\mathbf{D}_t\|_2 \leq \lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i. \tag{24}$$

1176 By Cauchy-Schwartz inequality and the triangle inequality, we can bound equation 22 by

$$\begin{aligned}
1177 & \left( \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right)^\top \left( \hat{\mathbf{A}}^{(t)} \right)^{\frac{1}{2}} \left( \hat{\mathbf{A}}^{(t)} \right)^{-\frac{1}{2}} \mathbf{D}_t \cdot \boldsymbol{\theta}_* \\
1178 & \leq \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}} \cdot \|\mathbf{D}_t\|_2 \cdot \|\boldsymbol{\theta}_*\|_{\left( \hat{\mathbf{A}}^{(t)} \right)^{-1}} \\
1179 & \\
1180 & \leq H \cdot \sqrt{\lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i} \cdot \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}, \\
1181 & \\
1182 & \\
1183 & \\
1184 & 
\end{aligned} \tag{25}$$

1185 where the last inequality holds because

$$1186 \quad \|\boldsymbol{\theta}_*\|_{\left( \hat{\mathbf{A}}^{(t)} \right)^{-1}}^2 \leq \frac{\|\boldsymbol{\theta}_*\|_2^2}{\lambda_{\min} \left( \hat{\mathbf{A}}^{(t)} \right)} \leq \frac{H^2}{\lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i}.$$

1188 **Bounding the variance error.** For the variance error, we have

$$\begin{aligned}
1189 & (\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_*)^\top \mathbf{X}_t^\top \mathbf{z}_t = (\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_*)^\top (\mathbf{A}^{(t)})^{-\frac{1}{2}} (\mathbf{A}^{(t)})^{\frac{1}{2}} \mathbf{X}_t^\top \mathbf{z}_t \\
1190 & \\
1191 & \\
1192 & \\
1193 & \leq \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}} \cdot \frac{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\mathbf{A}^{(t)}}}{\left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}} \cdot \left\| \mathbf{X}_t^\top \mathbf{z}_t \right\|_{(\mathbf{A}^{(t)})^{-1}} \\
1194 & \\
1195 & \\
1196 & \leq \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}} \cdot \left\| \mathbf{X}_t^\top \mathbf{z}_t \right\|_{(\mathbf{A}^{(t)})^{-1}}. \\
1197 & \\
1198 & 
\end{aligned} \tag{26}$$

1199 where the last inequality holds because for any vector  $\mathbf{a}$

$$\begin{aligned}
1200 & \\
1201 & \\
1202 & \|\mathbf{a}\|_{\mathbf{A}^{(t)}}^2 - \|\mathbf{a}\|_{\hat{\mathbf{A}}^{(t)}}^2 = \mathbf{a}^\top \left( \mathbf{X}^\top \mathbf{X} - \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)} - \sum_{i=1}^{B_t} \bar{\sigma}_i \mathbf{I} \right) \mathbf{a} \\
1203 & \\
1204 & = \mathbf{a}^\top \left( \mathbf{X}^\top \mathbf{X} - \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)} \right) \mathbf{a} - \sum_{i=1}^{B_t} \bar{\sigma}_i \|\mathbf{a}\|_2^2 \\
1205 & \\
1206 & \leq \sum_{i=1}^{B_t} \bar{\sigma}_i \|\mathbf{a}\|_2^2 - \sum_{i=1}^{B_t} \bar{\sigma}_i \|\mathbf{a}\|_2^2 \\
1207 & \\
1208 & = 0 \\
1209 & \\
1210 & \\
1211 & 
\end{aligned} \tag{27}$$

1212 By Proposition 1, we can bound the variance error term as follows

$$\begin{aligned}
1213 & \\
1214 & \\
1215 & (\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_*)^\top \mathbf{X}_t^\top \mathbf{z}_t \\
1216 & = (\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_*)^\top (\mathbf{A}^{(t)})^{-\frac{1}{2}} (\mathbf{A}^{(t)})^{\frac{1}{2}} \mathbf{X}_t^\top \mathbf{z}_t \\
1217 & \\
1218 & \leq \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}} \cdot \left\| \mathbf{X}_t^\top \mathbf{z}_t \right\|_{(\mathbf{A}^{(t)})^{-1}} \\
1219 & \\
1220 & \leq \sqrt{2R^2 \ln \left( \frac{1}{\delta} |\mathbf{A}^{(t)}|^{\frac{1}{2}} |\lambda \mathbf{I}|^{-\frac{1}{2}} \right)} \cdot \left\| \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_* \right\|_{\hat{\mathbf{A}}^{(t)}}. \\
1221 & \\
1222 & \\
1223 & 
\end{aligned} \tag{28}$$

1224 According to equation 27, we have  $|\hat{\mathbf{A}}^{(t)}| \geq |\mathbf{A}^{(t)}|$ . For any  $t \in [T]$ , since the rank of  $\hat{\mathbf{A}}^{(t)}$  is at  
1225 most  $2l_{B_t}$ , we can bound the determinant of  $\hat{\mathbf{A}}^{(t)}$  as follows

$$\begin{aligned}
1226 & \\
1227 & \\
1228 & |\hat{\mathbf{A}}^{(t)}| \leq \left( \sum_{i=1}^{B_t} \alpha_i^{(t)} + \lambda \right)^{d-2l_{B_t}} \cdot \prod_{i=1}^{2l_{B_t}} \lambda_i \left( \hat{\mathbf{A}}^{(t)} \right) \\
1229 & \\
1230 & \leq \left( \sum_{i=1}^{B_t} \alpha_i^{(t)} + \lambda \right)^{d-2l_{B_t}} \left( \frac{\sum_{i=1}^{2l_{B_t}} \lambda_i \left( \hat{\mathbf{A}}^{(t)} \right)}{2l_{B_t}} \right)^{2l_{B_t}} \\
1231 & \\
1232 & = \left( \sum_{i=1}^{B_t} \bar{\sigma}_i + \lambda \right)^{d-2l_{B_t}} \left[ \sum_{i=1}^{B_t} \bar{\sigma}_i + \lambda + \frac{\text{Tr} \left( \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)} \right)}{2l_{B_t}} \right]^{2l_{B_t}} \\
1233 & \\
1234 & \leq \left( \sum_{i=1}^{B_t} \bar{\sigma}_i + \lambda \right)^{d-2l_{B_t}} \left( \left( \sum_{i=1}^{B_t} \bar{\sigma}_i - \frac{\sum_{i=1}^{B_t} l_i \cdot \bar{\sigma}_i}{2l_{B_t}} \right) + \lambda + \frac{TL^2}{2l_{B_t}} \right)^{2l_{B_t}}, \\
1235 & \\
1236 & \\
1237 & \\
1238 & \\
1239 & \\
1240 & \\
1241 & 
\end{aligned} \tag{29}$$

where the last inequality satisfies due to

$$\begin{aligned}
\text{Tr} \left( \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)} \right) &= \sum_{i=1}^{B_t} \text{Tr} \left( (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)} \right) \\
&= \sum_{s=1}^t \text{Tr}(\mathbf{x}_s^\top \mathbf{x}_s) - \sum_{i=1}^{B_t} l_i \cdot \bar{\sigma}_i \\
&\leq TL^2 - \sum_{i=1}^{B_t} l_i \cdot \bar{\sigma}_i.
\end{aligned}$$

Therefore, the variance error term can be bounded as

$$\begin{aligned}
&(\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_\star)^\top \mathbf{X}_t^\top \mathbf{z}_t \\
&\leq \sqrt{2R^2 \ln \left( \frac{1}{\delta} |\mathbf{A}^{(t)}|^{\frac{1}{2}} |\lambda \mathbf{I}|^{-\frac{1}{2}} \right)} \cdot \|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_\star\|_{\hat{\mathbf{A}}^{(t)}} \\
&\leq \sqrt{2R^2 \ln \left( \frac{1}{\delta} |\hat{\mathbf{A}}^{(t)}|^{\frac{1}{2}} |\lambda \mathbf{I}|^{-\frac{1}{2}} \right)} \cdot \|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_\star\|_{\hat{\mathbf{A}}^{(t)}} \\
&\leq R \cdot \sqrt{2 \ln \left( \frac{1}{\delta} \right) + (d - 2l_{B_t}) \ln \left( 1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_t} \cdot \ln \left( 1 + \frac{tL^2}{2l_{B_t}\lambda} + \frac{h_t}{\lambda} \right)} \cdot \|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_\star\|_{\hat{\mathbf{A}}^{(t)}} \\
&\leq R \cdot \sqrt{2 \ln \left( \frac{1}{\delta} \right) + d \ln \left( 1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_t} \cdot \ln \left( 1 + \frac{tL^2}{2l_{B_t}\lambda} + \frac{h_t}{\lambda} \right)} \cdot \|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_\star\|_{\hat{\mathbf{A}}^{(t)}},
\end{aligned}$$

where  $h_t = \sum_{i=1}^{B_t} \bar{\sigma}_i - \frac{\sum_{i=1}^{B_t} l_i \cdot \bar{\sigma}_i}{2l_{B_t}}$ .

Sum up the bias error term and the variance error term and divide both sides by  $\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_\star\|_{\hat{\mathbf{A}}^{(t)}}$  simultaneously, we have

$$\begin{aligned}
\|\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}_\star\|_{\hat{\mathbf{A}}^{(t)}} &\leq R \cdot \sqrt{2 \ln \left( \frac{1}{\delta} \right) + d \ln \left( 1 + \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_t} \cdot \ln \left( 1 + \frac{tL^2}{2l_{B_t}\lambda} + \frac{h_t}{\lambda} \right)} \\
&\quad + H \cdot \sqrt{\lambda + \sum_{i=1}^{B_t} \bar{\sigma}_i},
\end{aligned}$$

which concludes the proof.  $\square$

Next, we start to prove the regret. Similar to the case using FD, since the algorithm uses the principle of optimism in the face of uncertainty to select the arm, we can bound instantaneous regret by



equation 19. Utilizing equation 19 and Cauchy-Schwartz inequality, we derive the following bound

$$\begin{aligned}
\text{Regret}_T &= \sum_{t=1}^T \max_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\top \boldsymbol{\theta}_* - \sum_{t=1}^T \mathbf{x}_t^\top \boldsymbol{\theta}_* \\
&\leq 2 \sum_{t=1}^T \min \left\{ HL, \hat{\beta}_{t-1}(\delta) \cdot \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}} \right\} \\
&\leq 2 \sum_{t=1}^T \hat{\beta}_{t-1}(\delta) \min \left\{ \frac{L}{\sqrt{\lambda}}, \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}} \right\} \\
&\leq 2 \cdot \max \left\{ 1, \frac{L}{\sqrt{\lambda}} \right\} \cdot \hat{\beta}_T(\delta) \cdot \sum_{t=1}^T \min \left\{ 1, \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}} \right\} \\
&\leq 2 \cdot \max \left\{ 1, \frac{L}{\sqrt{\lambda}} \right\} \cdot \hat{\beta}_T(\delta) \cdot \sqrt{T \sum_{t=1}^T \min \left\{ 1, \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}}^2 \right\}}.
\end{aligned} \tag{30}$$

We present a lemma of RFD-sketched leverage scores to conclude the proof.

**Lemma 3.**

$$\sum_{t=1}^T \min \left\{ 1, \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}}^2 \right\} \leq 2l_{B_T} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} + \frac{h_T}{\lambda} \right).$$

*Proof.* Denote  $\mathbf{C}_t = \hat{\mathbf{A}}^{(t-1)} + \mathbf{x}_t^\top \mathbf{x}_t$ . Notice that the first  $2l_{B_t}$  eigenvalues of  $\mathbf{C}_t$  are the same as  $\hat{\mathbf{A}}^{(t)}$  while the other eigenvalues of  $\mathbf{C}_t$  are  $\sum_{i=1}^{B_t} \alpha_i^{(t-1)} + \lambda$ . Thus we can obtain  $\frac{|\hat{\mathbf{A}}^{(t)}|}{|\mathbf{C}_t|} = \left( \frac{\sum_{i=1}^{B_t} \alpha_i^{(t)} + \lambda}{\sum_{i=1}^{B_{t-1}} \alpha_i^{(t-1)} + \lambda} \right)^{d-2l_{B_t}}$ .

For the determinant of  $\hat{\mathbf{A}}^{(t)}$ , we have

$$\begin{aligned}
|\hat{\mathbf{A}}^{(t)}| &= \left( \frac{\sum_{i=1}^{B_t} \alpha_i^{(t)} + \lambda}{\sum_{i=1}^{B_{t-1}} \alpha_i^{(t-1)} + \lambda} \right)^{d-2l_{B_t}} \cdot |\mathbf{C}_t| \\
&= \left( \frac{\sum_{i=1}^{B_t} \alpha_i^{(t)} + \lambda}{\sum_{i=1}^{B_{t-1}} \alpha_i^{(t-1)} + \lambda} \right)^{d-2l_{B_t}} \cdot |\hat{\mathbf{A}}^{(t-1)}| \cdot \left| \mathbf{I} + \left( \hat{\mathbf{A}}^{(t-1)} \right)^{-1} \mathbf{x}_t^\top \mathbf{x}_t \right| \\
&= \left( \frac{\sum_{i=1}^{B_t} \alpha_i^{(t)} + \lambda}{\sum_{i=1}^{B_{t-1}} \alpha_i^{(t-1)} + \lambda} \right)^{d-2l_{B_t}} \cdot |\hat{\mathbf{A}}^{(t-1)}| \cdot \left( 1 + \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}}^2 \right) \\
&= \left( \frac{\sum_{i=1}^{B_t} \bar{\sigma}_i + \lambda}{\lambda} \right)^{d-2l_{B_t}} \cdot |\lambda \mathbf{I}| \cdot \prod_{s=1}^t \left( 1 + \|\mathbf{x}_s\|_{(\hat{\mathbf{A}}^{(s-1)})^{-1}}^2 \right).
\end{aligned} \tag{31}$$

Since  $\min(1, x) \leq 2 \ln(1+x)$  for all  $x \geq 0$ , using equation 31, we can derive the following bound

$$\begin{aligned}
&\sum_{t=1}^T \min \left\{ 1, \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}}^2 \right\} \\
&\leq 2 \sum_{t=1}^T \ln \left( 1 + \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}}^2 \right) \\
&= 2 \cdot \ln \left( \left( \frac{\lambda}{\sum_{i=1}^{B_T} \bar{\sigma}_i + \lambda} \right)^{d-2l_{B_T}} \cdot \frac{|\hat{\mathbf{A}}^{(T)}|}{|\lambda \mathbf{I}|} \right) \\
&\leq 2l_{B_T} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} + \frac{h_T}{\lambda} \right),
\end{aligned}$$

where the last step holds by equation 29 and  $h_T = \sum_{i=1}^{B_T} \bar{\sigma}_i - \frac{\sum_{i=1}^{B_T} l_i \cdot \bar{\sigma}_i}{2l_{B_T}}$ .  $\square$

We combine equation 30, Theorem 6 and Lemma 3. Assume  $L \geq \sqrt{\lambda}$ , we have

$$\begin{aligned} \text{Regret}_T &= \sum_{t=1}^T \max_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\top \boldsymbol{\theta}_* - \sum_{t=1}^T \mathbf{x}_t^\top \boldsymbol{\theta}_* \\ &\leq 2 \cdot \max \left\{ 1, \frac{L}{\sqrt{\lambda}} \right\} \cdot \hat{\beta}_T(\delta) \cdot \sqrt{T \sum_{t=1}^T \min \left\{ 1, \|\mathbf{x}_t\|_{(\hat{\mathbf{A}}^{(t-1)})^{-1}}^2 \right\}} \\ &\stackrel{\hat{\sigma}}{=} \frac{L}{\sqrt{\lambda}} \cdot \sqrt{T} \cdot \sqrt{2l_{B_T} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} + \frac{h_T}{\lambda} \right)} \cdot \left( H \cdot \sqrt{\lambda + \sum_{i=1}^{B_T} \bar{\sigma}_i} + \right. \\ &\quad \left. R \cdot \sqrt{d \ln \left( 1 + \frac{\sum_{i=1}^{B_T} \bar{\sigma}_i}{\lambda} \right) + 2l_{B_T} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} + \frac{h_T}{\lambda} \right)} \right). \end{aligned}$$

According to Theorem 2, we can bound the spectral error by error  $\epsilon$ , which is

$$\begin{aligned} \text{Regret}_T &\stackrel{\hat{\sigma}}{=} \frac{L}{\sqrt{\lambda}} \cdot \sqrt{\ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} + \frac{h_T}{\lambda} \right)} \cdot \left( H \cdot \sqrt{\lambda + \epsilon} + \right. \\ &\quad \left. R \cdot \sqrt{d \ln \left( 1 + \frac{\epsilon}{\lambda} \right) + 2l_{B_T} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} + \frac{h_T}{\lambda} \right)} \right) \\ &\stackrel{\hat{\sigma}}{=} \frac{L}{\sqrt{\lambda}} \cdot \sqrt{\ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} + \left( 1 - \frac{2 - B_T}{2^{B_T+1}} \right) \cdot \frac{\epsilon}{\lambda} \right)} \cdot \left( H \cdot \sqrt{\lambda + \epsilon} + \right. \\ &\quad \left. R \cdot \sqrt{d \ln \left( 1 + \frac{\epsilon}{\lambda} \right) + 2l_{B_T} \cdot \ln \left( 1 + \frac{TL^2}{2l_{B_T}\lambda} + \left( 1 - \frac{2 - B_T}{2^{B_T+1}} \right) \cdot \frac{\epsilon}{\lambda} \right)} \right), \end{aligned}$$

where the last inequality holds because

$$\begin{aligned} h_T &= \sum_{i=1}^{B_T} \bar{\sigma}_i - \frac{\sum_{i=1}^{B_T} l_i \cdot \bar{\sigma}_i}{2l_{B_T}} \\ &= \sum_{i=1}^{B_T} \left( 1 - \frac{2^{i-1}}{2^{B_T}} \right) \cdot \bar{\sigma}_i \\ &\leq \epsilon \cdot \sum_{i=1}^{B_T} \left( 1 - \frac{2^{i-1}}{2^{B_T}} \right) \cdot \frac{1}{2^i} = \left( 1 - \frac{2 - B_T}{2^{B_T+1}} \right) \cdot \epsilon \end{aligned}$$

## G PROPERTIES OF DYADIC BLOCK SKETCHING FOR RFD

In this section, we highlight two significant properties of Dyadic Block Sketching for RFD that elucidate why the regret bound of DBSLinUCB using RFD is improved. Although Robust Frequent Directions for ridge regression have been studied by Luo et al. (2019), their theory is limited to single deterministic streaming sketches. We demonstrate that the decomposability of multi-scale sketching does not alter the properties of RFD.

We begin with the positive definite monotonicity of Dyadic Block Sketching for RFD, which ensures that the sequence of approximation matrices is per-step optimal.

**Theorem 7** (Positive Definite Monotonicity). *At round  $t$ , denote that the Dyadic Block Sketching for RFD provides a sketch  $\mathbf{S}^{(t)}$ , we have the following equation*

$$\left( \mathbf{S}^{(t)} \right)^\top \mathbf{S}^{(t)} + \alpha^{(t)} \mathbf{I} \succeq \left( \mathbf{S}^{(t-1)} \right)^\top \mathbf{S}^{(t-1)} + \alpha^{(t-1)} \mathbf{I}.$$

1404 *Proof.* Notice that  $\alpha^{(t)}\mathbf{I} + (\mathbf{S}^{(t)})^\top \mathbf{S}^{(t)} = \sum_{i=1}^{B_t} \alpha_i^{(t)}\mathbf{I} + \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)}$ , where  $\mathbf{S}_i^{(t)}$  is the  
 1405 sketch matrix in block  $i$  and  $\alpha_i^{(t)}$  is the adaptive regularization term of RFD at round  $t$ .  
 1406

1407  
 1408 Let  $\mathbf{Q} = \left[ (\mathbf{S}_{B_t}^{(t-1)})^\top, \mathbf{x}_t^\top \right]^\top$ ,  $\sigma_t$  is the shrinking singular values of active block at round  $t$ , the  
 1409 shrinking step of RFD provides  
 1410

$$1411 \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)} + \sigma_t \mathbf{I} \succeq \sum_{i=1}^{B_{t-1}} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)} + \mathbf{Q}^\top \mathbf{Q} \succeq \sum_{i=1}^{B_{t-1}} (\mathbf{S}_i^{(t-1)})^\top \mathbf{S}_i^{(t-1)}. \quad (32)$$

1412  
 1413 Therefore, for any unit vector  $\mathbf{a}$ , we have  
 1414

$$1415 \begin{aligned} & \mathbf{a}^\top \left( (\mathbf{S}^{(t)})^\top \mathbf{S}^{(t)} + \alpha^{(t)}\mathbf{I} - (\mathbf{S}^{(t-1)})^\top \mathbf{S}^{(t-1)} + \alpha^{(t-1)}\mathbf{I} \right) \mathbf{a} \\ &= \mathbf{a}^\top \left( \sum_{i=1}^{B_t} \alpha_i^{(t)}\mathbf{I} + \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)} - \sum_{i=1}^{B_{t-1}} \alpha_i^{(t-1)}\mathbf{I} - \sum_{i=1}^{B_{t-1}} (\mathbf{S}_i^{(t-1)})^\top \mathbf{S}_i^{(t-1)} \right) \mathbf{a} \\ &= \mathbf{a}^\top \left( \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)} + \sigma_t \mathbf{I} - \sum_{i=1}^{B_{t-1}} (\mathbf{S}_i^{(t-1)})^\top \mathbf{S}_i^{(t-1)} \right) \mathbf{a} \\ &\geq 0, \end{aligned}$$

1416  
 1417 which concludes the proof.  $\square$   
 1418  
 1419  
 1420  
 1421  
 1422  
 1423  
 1424  
 1425  
 1426  
 1427

1428  
 1429 Next, we prove that the sketch matrix produced by Dyadic Block Sketching for RFD is better  
 1430 conditioned than those produced by Dyadic Block Sketching for FD and the covariance matrix. In  
 1431 this context, the  $\alpha$  selected by RFD is optimal, as choosing a smaller  $\alpha$  would result in a worse  
 1432 condition number for the approximation matrices.  
 1433  
 1434

1435 **Theorem 8** (Well-Conditioned Property). *Let  $\text{cond}(\mathbf{X}) = \frac{\sigma_{\max}(\mathbf{X})}{\sigma_{\min}(\mathbf{X})}$  be the condition number of*  
 1436 *matrix  $\mathbf{X}$ . At round  $t$ , denote that the Dyadic Block Sketching for RFD provides a sketch  $\mathbf{S}^{(t)}$ , we*  
 1437 *have*  
 1438

$$1439 \text{cond} \left( (\mathbf{S}^{(t)})^\top \mathbf{S}^{(t)} + \alpha^{(t)}\mathbf{I} + \lambda \mathbf{I} \right) \leq \text{cond} \left( (\mathbf{S}^{(t)})^\top \mathbf{S}^{(t)} + \lambda \mathbf{I} \right),$$

$$1440 \text{cond} \left( (\mathbf{S}^{(t)})^\top \mathbf{S}^{(t)} + \alpha^{(t)}\mathbf{I} + \lambda \mathbf{I} \right) \leq \text{cond} (\mathbf{X}_t^\top \mathbf{X}_t + \lambda \mathbf{I}).$$

1441  
 1442 *Proof.* Notice that  $\alpha^{(t)}\mathbf{I} + (\mathbf{S}^{(t)})^\top \mathbf{S}^{(t)} = \sum_{i=1}^{B_t} \alpha_i^{(t)}\mathbf{I} + \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)}$ , where  $\mathbf{S}_i^{(t)}$  is the  
 1443 sketch matrix in block  $i$  and  $\alpha_i^{(t)}$  is the adaptive regularization term of RFD at round  $t$ . We have  
 1444  
 1445

$$1446 \text{cond} \left( (\mathbf{S}^{(t)})^\top \mathbf{S}^{(t)} + \alpha^{(t)}\mathbf{I} + \lambda \mathbf{I} \right) = \frac{\sigma_{\max} \left( \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)} \right) + \lambda + \sum_{i=1}^{B_t} \alpha_i^{(t)}}{\lambda + \sum_{i=1}^{B_t} \alpha_i^{(t)}}$$

$$1447 \leq \frac{\sigma_{\max} \left( \sum_{i=1}^{B_t} (\mathbf{S}_i^{(t)})^\top \mathbf{S}_i^{(t)} \right) + \lambda}{\lambda}$$

$$1448 = \text{cond} \left( (\mathbf{S}^{(t)})^\top \mathbf{S}^{(t)} + \lambda \mathbf{I} \right).$$

Similarly, we have

$$\begin{aligned} \text{cond} \left( \left( \mathbf{S}^{(t)} \right)^\top \mathbf{S}^{(t)} + \alpha^{(t)} \mathbf{I} + \lambda \mathbf{I} \right) &= \frac{\sigma_{\max} \left( \sum_{i=1}^{B_t} \left( \mathbf{S}_i^{(t)} \right)^\top \mathbf{S}_i^{(t)} \right) + \lambda + \sum_{i=1}^{B_t} \alpha_i^{(t)}}{\lambda + \sum_{i=1}^{B_t} \alpha_i^{(t)}} \\ &\leq \frac{\sigma_{\max} \left( \mathbf{X}_t^\top \mathbf{X}_t \right) + \lambda + \sum_{i=1}^{B_t} \alpha_i^{(t)}}{\lambda + \sum_{i=1}^{B_t} \alpha_i^{(t)}} \\ &\leq \frac{\sigma_{\max} \left( \mathbf{X}_t^\top \mathbf{X}_t \right) + \lambda}{\lambda} \\ &\leq \text{cond} \left( \mathbf{X}_t^\top \mathbf{X}_t + \lambda \mathbf{I} \right), \end{aligned}$$

which concludes the proof.  $\square$

## H OPEN ACCESS TO DATA

MNIST: <http://yann.lecun.com/exdb/mnist/>

## I EXPERIMENT OF ERROR PARAMETER

We evaluate the performance of the proposed method by varying the error parameter  $\epsilon$  on the real-world dataset MNIST. The experimental setup is consistent with Section 5.3. Specifically, we set  $\epsilon = 200, 1000, \text{ and } 8000$ , and record the spectral norm error  $\|\mathbf{A}_t^\top \mathbf{A}_t - \mathbf{S}_t^\top \mathbf{S}_t\|_2$ , regret, and running time.

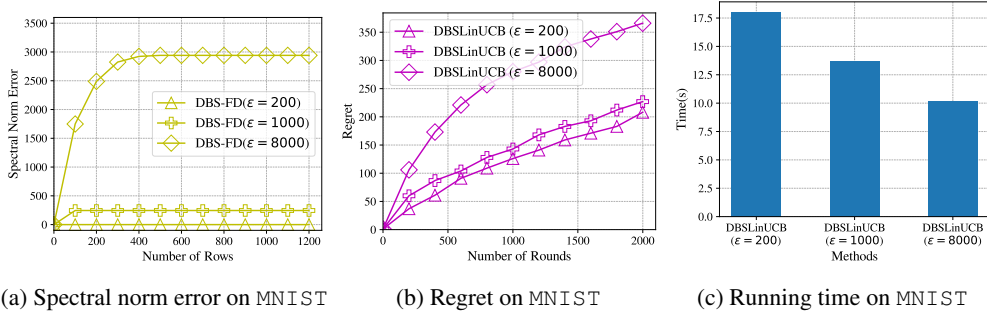


Figure 5: (a), (b), (c): The spectral norm error, cumulative regret and total running time w.r.t the error parameter  $\epsilon$  on MNIST

From Figure 5, we observe that increasing the error parameter  $\epsilon$  results in higher spectral error and regret but reduces computational time, which aligns with the theoretical results. Furthermore, in practical applications, it is unnecessary to set  $\epsilon$  too small. As shown in Figures 5a and 5b, the performance with  $\epsilon = 200$  is comparable to that with  $\epsilon = 1000$ . This is because  $\epsilon$  serves as an upper bound on the spectral error, and appropriate sketching operations have minimal impact on overall performance.