

Adaptive estimation of random vectors with bandit feedback: A mean-squared error viewpoint

Dipayan Sen¹, Prashanth L.A.¹ and Aditya Gopalan²

Abstract— We consider the problem of sequentially learning to estimate, in the mean squared error (MSE) sense, a Gaussian K -vector of unknown covariance by observing only $m < K$ of its entries in each round. We first establish a concentration bound for MSE estimation. We then frame the estimation problem with bandit feedback and propose a variant of the successive elimination algorithm. We also derive a minimax lower bound to understand the fundamental limit on the sample complexity of this problem.

I. INTRODUCTION

Several real-world applications involve collecting local measurements of a physical phenomenon, and then using the underlying correlation structure to form an estimate of the physical phenomenon over a wider region. For instance, using sensors to (i) monitor the temperature over a region [1] and (ii) detect contamination in a water distribution network [2]. Traffic monitoring in a cellular network is another application [3], where the underlying correlation structure plays a major role. To elaborate, the aim here is to collect traffic load measurements from a handful of base stations to form an estimate of the load on all base stations.

We assume that the underlying distribution is K -dimensional Gaussian with covariance matrix Σ : a model that has been shown to be practically viable in [3]. We employ the mean squared error (MSE) objective to capture the underlying correlation structure. For a m -subset A , the MSE $\psi(A)$ is given by

$$\psi(A) = \text{Tr} \left(\Sigma_{A'A'} - \Sigma_{A'A} (\Sigma_{AA})^{-1} \Sigma_{AA'} \right), \quad (1)$$

where A' is $[K] \setminus A$, and $\Sigma_{AA}, \Sigma_{A'A'}, \Sigma_{A'A}, \Sigma_{AA'}$ are sub-matrices of Σ in obvious notation (See Section II for the details). We first consider the problem of estimating the MSE of a m -subset, say A , given a batch of i.i.d. samples for each of the sub-matrices listed above. This problem is non-adaptive in the sense that each sub-matrix entry is pulled equally. An adaptive version of this problem is when we are provided entry-wise estimates of Σ , with non-uniform sampling. Such a set of samples facilitates estimation of MSE of any m -subset A .

From a statistical learning viewpoint, significant progress has been made on the problem of covariance matrix attention (cf. [4]). However, the problem of MSE estimation has not received enough attention, and there are no concentration

bounds available for the problem of estimating (1), to the best of our knowledge. We propose a natural MSE estimator based on sample-averages for the non-adaptive as well as the adaptive settings. Since the sample average estimator of Σ_{AA} may not be invertible, we perform an eigen-decomposition followed by projection of eigenvalues to the positive side. Next, we derive concentration bounds for the MSE estimation problem in the non-adaptive and adaptive settings. The bounds that we derive exhibit an exponential tail decay in either case.

We then frame the adaptive estimation problem with bandit feedback in the best-arm identification framework setting [5]. We apply the successive elimination technique [6] to cater to the adaptive estimation problem. We present an upper bound on the sample complexity of this algorithm. Further, to understand the fundamental limit on the sample complexity of this adaptive estimation bandit problem, we derive an information-theoretic lower bound. We construct a set of covariance matrices that are rich enough to include the least favorable instance for any bandit algorithm. We establish the lower bound using the well-known standard change of measure argument by constructing problem transformations based on the aforementioned set of covariance matrices, but the technical steps require significant deviations in terms of algebraic effort. Moreover, the setting we consider involve sampling more than one arm, which is strictly necessary for estimating the underlying correlation. This sampling change implies additional effort in computing certain KL-divergences, which are then related to the sub-optimality gap in MSEs.

Related work. Previous works such as [7], [8], [9] feature bandit formulations where the underlying correlation structure appears in the objective. In [7], the aim is to find the maximum correlated subset, i.e., a set that has highly correlated members. In contrast, our goal is to find a subset that best captures information about other, as quantified by the MSE objective. In addition, unlike [7], we do not assume unit variances in the underlying model. Next, in [8], which is the closest related work, the authors propose an MSE-based objective for a simplified version of the problem where the goal is to find an arm (or 1-subset) that is most correlated to the remaining $K - 1$ arms in the MSE sense. Our problem formulation is more general as we consider MSE of m -subsets, with $1 \leq m \leq K$. This generalization leads to bigger technical challenges in MSE estimation and concentration, as well as in the lower bound analysis. Finally, in [9], the authors assume that the arms are correlated through a latent random source, and the objective

¹Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai 600036, India, {cs18s012, prashla}@cse.iitm.ac.in

²Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru 560012, India, aditya@iisc.ac.in

is to identify the arm with the highest mean. In [10], the authors study the impact of correlation on the regret, while featuring a regular bandit formulation, i.e.. of identifying the arm with the highest mean.

The rest of the paper is organized as follows: In Section II, we formally define the notion of MSE. In Sections III and IV, we describe MSE estimation in the non-adaptive and adaptive settings, respectively. In Section V, we formulate the adaptive MSE estimation problem with bandit feedback, and we present a variant of successive elimination algorithm for solving this problem. In Section VI, we present a minimax lower bound on the sample complexity of the adaptive estimation problem in a BAI framework. Due to space limitations, we provide detailed proofs in the longer version of the paper, which is available in [11]. Finally, in Section VII we provide our concluding remarks.

II. PRELIMINARIES

We consider a jointly Gaussian K -vector $X = (X_1, \dots, X_K)$, with mean zero and covariance matrix $\Sigma \triangleq \mathbb{E}[X^T X]$:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \dots & \rho_{1K}\sigma_1\sigma_K \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \dots & \rho_{2K}\sigma_2\sigma_K \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1K}\sigma_1\sigma_K & \rho_{2K}\sigma_2\sigma_K & \dots & \sigma_K^2 \end{bmatrix}, \quad (2)$$

where σ_i^2 , $i \in [K]$ is the variance of arm i and ρ_{ij} , $i, j = 1, \dots, K$, $i \neq j$, the correlation coefficient between arms i and j . Here $[n] = \{1, \dots, n\}$, for any natural number n .

Let \mathcal{A} denote the set of subsets of $[K]$ with cardinality m . The mean-squared error (MSE) for a given subset $A = \{i_1, \dots, i_m\} \in \mathcal{A}$ is defined as

$$\psi(A) \triangleq \sum_{j=1}^K \mathbb{E}[(X_j - \mathbb{E}[X_j | X_{i_1}, \dots, X_{i_m}])^2]. \quad (3)$$

As shown in [12], the above definition is equivalent to

$$\psi(A) = \text{Tr} \left(\Sigma_{A'A'} - \Sigma_{A'A} (\Sigma_{AA})^{-1} \Sigma_{AA'} \right), \quad (4)$$

where Tr denotes the trace function, $A' = [K] \setminus A$ is the complement of A , $\Sigma_{A'A'}$ (resp. Σ_{AA}) is the covariance matrix, which is obtained by restricting Σ to the set A' (resp. A).

We consider three problems related to the estimation of the MSE defined above.

III. NON-ADAPTIVE ESTIMATION

To estimate $\psi(A)$, it is apparent from (4) that we require an estimate of the sub-matrices Σ_{AA} , $\Sigma_{A'A}$, $\Sigma_{AA'}$, and $\Sigma_{A'A'}$. In the non-adaptive setting, we are given i.i.d. samples for each of the sub-matrices Σ_{AA} , $\Sigma_{A'A}$, $\Sigma_{AA'}$, and $\Sigma_{A'A'}$, for a given subset A . Using these samples from the underlying multivariate Gaussian distribution, we form the sample covariance matrices $\widehat{\Sigma}_{AA}$, $\widehat{\Sigma}_{AA'}$, $\widehat{\Sigma}_{A'A}$, and $\widehat{\Sigma}_{A'A'}$ to estimate the aforementioned four sub-matrices.

The ‘sample-average’ estimator $\widehat{\Sigma}_{AA}$ is not guaranteed to be invertible (though it is positive definite with high probability), while MSE estimation requires an estimate of Σ_{AA}^{-1} . To handle invertibility, we form the matrix $\widehat{\Sigma}_{A'A'}^+$ by performing an eigen-decomposition of $\widehat{\Sigma}_{AA}$, followed by a projection of eigenvalues to the positive side. Formally, for $i = 1, \dots, m$, let $\hat{\lambda}_i$ denote the eigenvalue of $\widehat{\Sigma}_{AA}$, with corresponding eigenvector v_i . The estimator $\widehat{\Sigma}_{AA}^+$ is defined by

$$\widehat{\Sigma}_{AA}^+ \triangleq \sum_{i=1}^m \hat{\lambda}_i^+ v_i v_i^T, \quad (5)$$

$$\text{where } \hat{\lambda}_i^+ = \begin{cases} \hat{\lambda}_i & \text{if } |\hat{\lambda}_i| \geq \zeta, \\ \zeta & \text{otherwise,} \end{cases}$$

for $i = 1, \dots, m$. It is easy to see that $\widehat{\Sigma}_{AA}^+$ is positive definite.

The MSE $\psi(A)$ associated with set A is then estimated as follows:

$$\widehat{\psi}(A) \triangleq \text{Tr} \left(\widehat{\Sigma}_{A'A'} - \widehat{\Sigma}_{A'A} (\widehat{\Sigma}_{AA}^+)^{-1} \widehat{\Sigma}_{AA'} \right). \quad (6)$$

Next, we proceed to analyze the concentration properties of the estimator defined above. For the sake of analysis, we make the following assumptions:

(A1). $0 < l = \min_i \sigma_i^2, \sigma_i^2 \leq 1$ for $i = 1, \dots, K$.

(A2). $\max(\|\Sigma_{AA}\|_2, \|\Sigma_{A'A'}\|_2) \leq M_0$, and $\|\Sigma_{AA}^{-1}\|_2 \leq \frac{1}{M_1}$, where $\|\cdot\|_2$ is the operator norm.

Assumption (A1) is used for the simpler 1-subset MSE estimation in [8], while (A2) is common in the analysis of covariance matrix estimates (cf. [13]). We now present a concentration bound for the MSE estimator (6).

Proposition 1 (MSE concentration: Non-adaptive case). Assume (A1) and (A2). Let $n_{AA}, n_{A'A}, n_{AA'}, n_{A'A'}$ denote the number of samples used to form $\widehat{\Sigma}_{AA}, \widehat{\Sigma}_{A'A}, \widehat{\Sigma}_{AA'},$ and $\widehat{\Sigma}_{A'A'}$, respectively.

Set the projection parameter ζ in (5) as $\zeta = M_0 \min \left(\sqrt{\frac{m+\log(\frac{1}{\delta})}{n_{AA}}}, \frac{m+\log(\frac{1}{\delta})}{n_{AA}} \right)$ and $n' = \min(n_{AA}, n_{A'A'}, n_{AA'}, n_{A'A})$.

Then, for any $0 < \epsilon < \eta \triangleq \min(2K, \lambda_{\min}(\Sigma_{AA}))^1$, the MSE estimate $\widehat{\psi}(A)$ defined by (6) satisfies

$$\mathbb{P} \left(|\widehat{\psi}(A) - \psi(A)| \geq \epsilon \right) \leq \underbrace{C_0 \exp \left[\frac{-n'}{mK^2(1+\eta)^3} \min \left[\frac{\epsilon}{12G_0}, \frac{\epsilon^2}{G_0^2} \right] \right]}_{(I)} + \underbrace{2mK \exp \left[-\frac{n'\epsilon^2}{72 C_3^2 m^2 K^4 (1+\eta)^3} \right]}_{(II)}, \quad (7)$$

¹ $\lambda_{\min}(AA)$ denotes the smallest eigenvalue of the matrix AA .

where $c = \frac{1}{[\lambda_{\min}(\Sigma_{AA}) - \epsilon]}$, $C_0 = [13mK + e^{m+K}]$,
 $C_1 = \frac{108\sqrt{2}}{l} \left(c + \frac{1}{M_1}\right)$, $C_2 = \left(\frac{3cM_0}{M_1}\right)$, $C_3 = \left(c + \frac{1}{M_1}\right)$,
and

$$G_0 = \max \left(\sqrt{m(K-m)^2(1+\eta)} C_1, \right. \\ \left. \sqrt{m(K-m)^2(1+\eta)} C_2, 3(K-m)M_0 \right).$$

Proof. The proof is available in Appendix A of [11]. \square

In the result above, we have $\epsilon \leq 2K$, and this constraint is not restrictive since the MSE $\psi(A) \leq K$ for any subset A in lieu of (A1). A similar observation holds for the adaptive case handled later.

To understand the terms (I) and (II) in (7), we have to look at the following decomposition of the MSE estimation error:

$$\begin{aligned} & \widehat{\psi}(A) - \psi(A) \\ &= \text{Tr} \left(\left(\widehat{\Sigma}_{A'A'} - \Sigma_{A'A'} \right) - \left[\widehat{\Sigma}_{A'A} \left(\widehat{\Sigma}_{AA}^+ \right)^{-1} \times \right. \right. \\ & \quad \left. \left(\widehat{\Sigma}_{AA'} - \Sigma_{AA'} \right) + \widehat{\Sigma}_{A'A} \left[\left(\widehat{\Sigma}_{AA}^+ \right)^{-1} - \Sigma_{AA}^{-1} \right] \Sigma_{AA'} \right. \\ & \quad \left. + \left(\widehat{\Sigma}_{A'A} - \Sigma_{A'A} \right) \Sigma_{AA}^{-1} \Sigma_{AA'} \right]. \end{aligned} \quad (8)$$

The first and third terms on the RHS above relate to estimation of a covariance matrix and its inverse. These terms lead to the term (I) in the bound (7) above. On the other hand the second and fourth terms on the RHS above relate to concentration of sample standard deviation and sample correlation coefficient, in turn leading to the term (II) in the bound (7).

IV. ADAPTIVE ESTIMATION

In the adaptive setting, we consider non-uniform sampling of the underlying covariance matrix, with the aim of reusing samples to estimate the MSE for different subsets.

The estimate in (6) is useful if one is concerned with estimating the MSE for a given subset. On the other hand, if one has to reuse sample information to estimate MSE for many subsets, then an approach that could be adopted is to maintain an estimate of each entry of the covariance matrix, and then, form MSE estimates for any subset by extracting the relevant information from the sample covariance matrix. We present a MSE estimation scheme based on this approach below.

For a subset $A = \{i_1, \dots, i_m\}$, the MSE $\psi(A)$, given in (4), can be re-written as follows:

$$\psi(A) = \sum_{j=1}^K [\sigma_j^2 - C_j (\Sigma_{AA}^{-1}) C_j^T], \quad (9)$$

where Σ_{AA} is as defined before, and $C_j = [\rho_{j i_1} \sigma_{i_1} \sigma_j \dots \rho_{j i_m} \sigma_{i_m} \sigma_j]$. The MSE expressions in (3), (4) and (9) are equivalent. We have chosen to use

(9) for adaptive estimation as it can be related easily to the MSE estimate presented below. Notice that, unlike the non-adaptive setting, the same sample set here can be used to estimate the MSE of any subset $A \in \mathcal{A}$.

From (9), it is apparent that one requires an estimate of the underlying variances, and correlation coefficients. Formally, we are given n_i samples for the variance σ_i^2 , and n_{ij} samples for the correlation coefficient ρ_{ij} , $i, j \in [K], i \neq j$. The aim is to estimate (9) using these samples. For $j = 1, \dots, K$ and $k = 1, \dots, m$, let $\hat{\rho}_{j i_k}$ denote the sample correlation coefficient, and let $\hat{\sigma}_j^2$ denote the sample variance. These quantities are formed using n_j and $n_{j i_k}$ samples, respectively, as follows:

$$\hat{\sigma}_j^2 = \overline{X_j^2}, \hat{\rho}_{j i_k} = \frac{\overline{X_j X_{i_k}}}{\hat{\sigma}_j \hat{\sigma}_{i_k}}, \text{ where } \overline{X_j^2} = \frac{1}{n_j} \sum_{t=1}^{n_j} X_{jt}^2, \\ \overline{X_j X_{i_k}} = \frac{1}{n_{j i_k}} \sum_{t=1}^{n_{j i_k}} X_{jt} X_{i_k t}.$$

Using the sample variance and sample correlation coefficients, we estimate the MSE $\psi(A)$ as follows:

$$\widehat{\psi}(A) = \sum_{j=1}^K \left[\hat{\sigma}_j^2 - \widehat{C}_j \left(\widehat{\Sigma}_{AA}^+ \right)^{-1} \widehat{C}_j^T \right]. \quad (10)$$

where $\widehat{C}_j = [\hat{\rho}_{j i_1} \hat{\sigma}_{i_1} \hat{\sigma}_j \dots \hat{\rho}_{j i_m} \hat{\sigma}_{i_m} \hat{\sigma}_j]$, $\widehat{\Sigma}_{AA}$ formed by using the relevant sample correlation coefficients $\hat{\rho}_{i_k i_l}, i_k, i_l \in A$, and sample variances $\hat{\sigma}_{i_k}^2, i_k \in A$, and $\widehat{\Sigma}_{AA}^+$ is defined in (5).

Under the assumptions that are identical to the non-adaptive setting, we present a concentration bound for the MSE estimator (10) in the result below.

Proposition 2 (MSE concentration: Adaptive case). *Assume (A1) and (A2). Set the projection parameter ζ in (5) as follows:*

$$\zeta = \sqrt{\frac{(1+\eta)^3(m^2-m)}{n''l^2}} \sqrt{\log \left(\frac{15(m^2-m)}{\delta} \right)} \\ + \sqrt{\frac{m \log \left(\frac{m}{\delta} \right)}{n''}},$$

where $\eta \triangleq \min(2K, \lambda_{\min}(\Sigma_{AA}))$. Then, for any $0 < \epsilon < \eta$, the MSE estimate formed using (10) satisfies

$$\begin{aligned} & \mathbb{P} \left(|\widehat{\psi}(A) - \psi(A)| \geq \epsilon \right) \leq \\ & 14mK \exp \left[-\frac{n''}{G_1} \min \left[\frac{l\epsilon}{12\sqrt{2} G_2}, \frac{l^2 \epsilon^2}{G_2^2 m(1+\eta)^2} \right] \right] \\ & + 30m^2 K \exp \left(-\frac{n''l^2 \epsilon^2}{G_3 (m^4 - m^2)(1+\eta)^7} \right), \end{aligned} \quad (11)$$

where $n'' = \min(n_i, n_j, n_{ij}, (i, j) \in [K], i \neq j)$, C_4 is a universal constant. $C_5 = 160 \left(c + \frac{1}{M_1}\right)$, $C_6 = \left(\frac{3c}{M_1}\right)$, $C_7 = \left(c + \frac{1}{M_1}\right)$, $G_1 = \max(8, m(1+\eta)^3)$, $G_2 = \max(1, C_5)$, and $G_3 = \max(C_4, C_4 C_6^2, 72 C_7^2)$.

Proof. The proof is available in Appendix B of [11]. \square

V. ADAPTIVE ESTIMATION WITH BANDIT FEEDBACK

We consider the fixed confidence variant of the best-arm identification framework [5]. In this setting, the interaction of a bandit algorithm with the environment is given below.

Adaptive estimation with bandit feedback

Input: set of m -subsets \mathcal{A} .

For all $t = 1, 2, \dots$, **repeat**

- 1) Select an m -subset $A_t \in \mathcal{A}$.
- 2) Observe a sample from the multi-variate Gaussian distribution corresponding to the arms in the set A_t .
- 3) Choose to continue, or stop and output an m -subset.

A subset that has the lowest MSE is considered optimal, i.e.,

$$A^* \in \arg \min_{A \in \mathcal{A}} \psi(A).$$

The aim in this setting is to devise an algorithm that outputs the best m -subset with high probability, while using a low number of samples. More precisely, for a given confidence parameter $\delta \in (0, 1)$, an algorithm is δ -PAC if it stops after τ rounds, and outputs a set A_τ that satisfies $\mathbb{P}(A_\tau \neq A^*) \leq \delta$. Among δ -PAC algorithms, the algorithm with minimum sample complexity $\mathbb{E}[\tau]$ is preferred.

For any set A , define

$$\Delta(A) \triangleq \psi(A) - \psi(A^*), \text{ and } \Delta = \min_{A \in \mathcal{A}} \Delta(A). \quad (12)$$

In the above, $\Delta(A)$ denotes the gap in MSE associated with a subset A , while Δ denotes the smallest gap. The upper and lower bounds that we derive subsequently features these quantities.

Successive Elimination For Correlated Bandits

In the fixed confidence setting that we consider, a naive algorithm based on Algorithm 1 in [6] would pull each subset equal number of times. Such an uniform sampling will be useful if all the subsets can capture the same amount of information about other subsets, i.e., when the underlying correlations and the variances are similar. However, with uneven correlations, uniform sampling does not make sense. The possible set of candidates for the most informative subset need to be sampled more than the other subsets in order to reduce the probability of error in identifying the best m -subset, and successive elimination [6] is an approach that embodies this idea.

We propose a variant of the successive elimination algorithm that is geared towards finding the best m -subset under the MSE objective. The algorithm maintains an active set, which is initialized to the set of all m -subsets \mathcal{A} . In each round t , the algorithm pulls each active m -subset once, and its MSE is estimated using (10). Following this, the algorithm eliminates all subsets whose confidence intervals are clearly separated

from the confidence interval of the empirically best subset seen so far, i.e., the one with the least MSE estimate. The algorithm terminates when there is only one m -subset left in the active set, and this event occurs with probability one.

For deriving the confidence width to be used in the successive elimination algorithm, we first re-write the bound derived in Proposition 2 as follows:

$$\begin{aligned} & \mathbb{P}\left(|\widehat{\psi}(A) - \psi(A)| \geq \epsilon\right) \\ & \leq K \left((13m + 1 + (30m^2 - 26m)) \exp\left(-\frac{n'' c_3 \epsilon^2}{c_1 + c_2 \epsilon}\right) \right), \end{aligned} \quad (13)$$

where $c_1 = \frac{G_1 G_2^2 m (1+\eta)^2}{l^2}$, $c_2 = \frac{12\sqrt{2} G_1 G_2}{l}$ and $c_3 = \frac{G_2 (m^4 - m^2) (1+\eta)^7}{l^2}$.

Inverting the tail bound in (13) leads to the following confidence width:

$$\alpha_t = \frac{c_2 \log \left[\frac{70 \binom{K}{m} K m^2 t^2}{\delta} \right]}{2c_3 t} + \sqrt{\frac{c_1 \log \left(\frac{70 \binom{K}{m} K m^2 t^2}{\delta} \right)}{2c_3 t}}. \quad (14)$$

The complete algorithm is given below.

Successive elimination for correlated bandits

Input: set of all m -subsets \mathcal{A} , $|\mathcal{A}| = \binom{K}{m}$, $\delta > 0$.

Initialization: set of active subsets $\mathcal{S} = \mathcal{A}$

For all $t = 1, 2, \dots$, **repeat**

- 1) Select all active m -subsets $A_t \in \mathcal{S}$.
- 2) Observe a sample from the m -variate Gaussian distribution corresponding to the arms in each of the active sets A_t .
- 3) Remove those subsets from \mathcal{S} such that $\widehat{\psi}_{A_t^*} - \widehat{\psi}_{A_t} \geq 2\alpha_t$, where α_t is defined in (14) and A_t^* is any active optimal subset at time t with minimum MSE, i.e., $A_t^* \in \arg \min_{A_t \in \mathcal{S}} \psi(A_t)$.
- 4) Continue until there is only one active m -subset in \mathcal{S} .

Fig. 1: Operational flow of successive elimination for correlated bandits.

We now present a bound on the sample complexity of the successive elimination algorithm for correlated bandits.

Theorem 1. Assume (A1) and (A2) for every $A \in \mathcal{A}$. The successive elimination algorithm is $(0, \delta)$ -PAC for any $\delta \in (0, 1)$, and w.p. at least $1 - \delta$, its sample complexity is bounded by

$$\mathcal{O} \left(\sum_{A \in \mathcal{A}} \frac{1}{\Delta(A)} \log \left(\frac{\binom{K}{m} K m^2 \log(\Delta(A)^{-1})}{\delta} \right) \right),$$

where $\Delta(A)$ is defined in (12).

Proof. The proof is available in Appendix C of [11]. \square

The sample complexity bound in the result above features the total number of m -subsets $\binom{K}{m}$, and is of the form $O\left(\frac{\binom{K}{m}}{\Delta} \log\left(\frac{\binom{K}{m}}{\delta}\right)\right)$, where Δ denotes the smallest gap. It is unclear if this bound can be improved without additional assumptions on the underlying covariance matrix Σ , and we believe the number of m -subsets $\binom{K}{m}$ has to appear in the sample complexity bound for a general covariance matrix Σ .

VI. LOWER BOUND

We consider a special case of the adaptive estimation problem, where the goal is to identify the best pair of arms, i.e.,

$$(i_1^*, i_2^*) \in \arg \min_{(i,j) \in [K] \times [K], i \neq j} \psi(\{i, j\}).$$

Let $Alg(\delta, K)$ denote the class of algorithms that are δ -PAC for the best pair identification problem. A lower bound on the sample complexity of this problem is presented below.

Theorem 2. *For any δ -PAC algorithm, there exists a bandit problem instance governed by a covariance matrix Σ such that the sample complexity $\mathbb{E}_\Sigma[\tau_\delta]$ of this algorithm satisfies*

$$\mathbb{E}_\Sigma[\tau_\delta] \geq \frac{\log\left(\frac{1}{2.4\delta}\right)}{\Delta}. \quad (15)$$

where Δ denotes the smallest gap on the problem instance governed by Σ .

Comparing the lower bound to the upper bound for successive elimination in Theorem 1, we observe that the dependence on the minimum gap Δ is the same in either bound. However, the lower bound does not have a dependency on m and K through the number of arms $\binom{K}{m}$ — a dependency that is present in the upper bound. We believe the lower bound is sub-optimal from the dependence on the number of m -subsets (or arms), and it would be an interesting future direction of future work to establish a lower bound that involves the $\binom{K}{m}$ factor.

The proof strategy is to use the following class of covariance matrices parameterized by ρ :

$$\Sigma = \begin{bmatrix} 1 & \rho & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho^2 & \rho^2 & \dots & \rho^2 \\ \rho & \rho^2 & 1 & \rho^3 & \dots & \rho^3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho^2 & \rho^3 & \dots & \rho^{K-1} & 1 \end{bmatrix} \quad (16)$$

Using Sylvester's criterion, it is easy to see that the matrix defined above is positive semi-definite.

For a K -armed Gaussian bandit instance with the underlying distribution governed by Σ defined above, the pair $\{1, 2\}$ has the least MSE.

We form $(2K - 4)$ transformations of the bandit instance described in (16). The transformations are achieved by relabelling the m^{th} row as either the first or second row of Σ , $m = 3, \dots, K$. Let us denote the pdf associated with

the original bandit instance by \mathcal{G} and \mathcal{G}^{km} is the probability density function (pdf) of the transformed bandit instance obtained by relabelling the k^{th} ($k = 1$ or 2) row and the m^{th} row of Σ .

The underlying covariance matrix for the problem instance corresponding to the m th transformation is Σ_{km} with m th row re-labelled as either row 1 or 2. Let $KL_{ij}^{km} \triangleq KL(\nu_i \nu_j || \nu'_i \nu'_j)$ specify the KL-divergence between $\nu_i \nu_j$ and $\nu'_i \nu'_j$, with the latter distribution derived from \mathcal{G}^{km} .

In the proof, we first show that

$$\begin{aligned} & \min_{\{w_{ij}\}} \mathbb{E}[\tau_\delta] \\ & \geq \frac{\log\left(\frac{1}{2.4\delta}\right)}{\max_{w \in \Delta_{\binom{K}{2}}} \min_{\Sigma' \in \text{Alt}(\Sigma)} \sum_{i,j} w_{ij} KL\left[\Sigma_{X_i X_j} || \Sigma'_{X_i X_j}\right]}, \quad (17) \end{aligned}$$

where $\Delta_{\binom{K}{2}}$ is the set of probability distributions on the arm-pairs, and $\text{Alt}(\Sigma) = \{\Sigma^{1m}, \Sigma^{2m}, m > 3\}$ is the set of transformed covariance matrices. While derivation of the inequality above is a straightforward variation to the proof in the classic bandit setting (cf. [14]), the rest of the proof in our case requires significant deviations. In particular, unlike the regular bandit case, the KL-divergences in the RHS above are not univariate. Moreover, deriving an upper bound on the max-min, which is defined in the RHS above, requires arguments that are specific to our correlated bandit setting.

We would like to note that the authors in [8] provide a lower bound for the correlated bandit problem with $m = 1$. The proof of the lower bound for the case of $m = 2$ is significantly different from the proof for $m = 1$. In particular, it is challenging since the proof involves KL-divergences for bivariate distributions and relating these KL-divergences to the underlying gaps involves tools from optimization (see the proof sketch below), as well as significant algebraic effort to simplify KL-divergence bounds inside the max-min in (17), and then, relating the simplified expression to the gap in MSEs of the original problem instance. Further, unlike [8], the ideas in our proof for $m = 2$ could be generalized to $m > 2$.

Proof. (Theorem 2) We provide a sketch of the proof here. The reader is referred to Appendix D of [11] for a detailed proof.

Notice that $\mathbb{E}_\Sigma[\tau_\delta] = \sum_{(i,j)} \mathbb{E}[N_{ij}(\tau_\delta)]$. For any $\Sigma' \in \text{Alt}(\Sigma)$, from Lemma 1 and Remark 2 of [14], we have

$$\sum_{(i,j)} \mathbb{E}[N_{ij}(\tau_\delta)] KL\left(\Sigma_{X_i X_j} || \Sigma'_{X_i X_j}\right) \geq \log\left(\frac{1}{2.4\delta}\right).$$

Consider the following optimization problem, with $\alpha_{ij} \triangleq \mathbb{E}[N_{ij}(\tau_\delta)]$:

$$\min_{\{\alpha_{ij}\}} \sum_{(i,j)} \alpha_{ij} \quad \text{s.t. for any } \Sigma' \in \text{Alt}(\Sigma)$$

$$\mathbb{E}[\tau_\delta] \sum_{(i,j)} \frac{\alpha_{ij}}{\mathbb{E}[\tau_\delta]} KL \left(\Sigma_{X_i X_j} \| \Sigma'_{X_i X_j} \right) \geq \log \left(\frac{1}{2.4\delta} \right).$$

Letting $w_{ij} \triangleq \frac{\alpha_{ij}}{\mathbb{E}[\tau_\delta]}$, the problem defined above is equivalent to the problem (17) defined earlier in Section VI.

Next, we sketch the derivation of an upper bound on the max-min in the denominator of (17).

Let $f(w, \Sigma') \triangleq \sum_{(i,j)} w_{ij} KL \left(\Sigma_{X_i X_j} \| \Sigma'_{X_i X_j} \right)$. Then,

$$\begin{aligned} f(w, \Sigma^{13}) &= w_{12} KL_{12}^{13} + w_{14} KL_{14}^{13} + \dots + w_{1K} KL_{1K}^{13} \\ &\quad + w_{23} KL_{23}^{13} + w_{34} KL_{34}^{13} + \dots + w_{3K} KL_{3K}^{13} \\ &\leq \frac{\rho^2}{2(1-\rho^2)} \times \\ &\quad \left(w_{12}(1-\rho^2) + (w_{14} + \dots + w_{1K})(1-\rho^4) \right. \\ &\quad \left. + w_{23}(1-\rho) + (w_{34} + \dots + w_{3K})(1-\rho^2) \right), \end{aligned}$$

where we applied a standard result for the KL-divergence between multivariate Gaussian distributions to bound KL_{ij}^{13} above. Along similar lines, we can bound $f(w, \Sigma^{1j}), j = 4, \dots, K$, and $f(w, \Sigma^{2j}), j = 3, \dots, K$.

Using these bounds, and letting

$\mathcal{H} = \{\alpha_1, \alpha_2, \dots, \alpha_K \geq 0, \alpha_1 + \alpha_2 + \dots + \alpha_K = 1\}$, we have

$$\min_w f(w, \Sigma^{1m}) \geq \max_w f(w, \Sigma^{2m}),$$

and

$$\begin{aligned} &\max_w \min (f(w, \Sigma^{23}), f(w, \Sigma^{24}), \dots, f(w, \Sigma^{2K})) \\ &= \max_w \min_{\mathcal{H}} \{ \alpha_1 f(w, \Sigma^{23}) + \dots + \alpha_K f(w, \Sigma^{2K}) \} \\ &\leq \min_{\mathcal{H}} \max_w \{ \alpha_1 f(w, \Sigma^{23}) + \dots + \alpha_K f(w, \Sigma^{2K}) \} \\ &\leq \max_w f(w, \Sigma^{23}) \quad (\text{choosing } \alpha_1 = 1, \alpha_i = 0, i \geq 2) \\ &= \max_w \left\{ \frac{\rho^4}{2(1-\rho^4)} \left((1-\rho^2)(w_{24} + \dots + w_{2K}) \right. \right. \\ &\quad \left. \left. + (1-\rho)(w_{34} + \dots + w_{3K}) \right) \right\} \\ &= \frac{\rho^4}{2(1+\rho^2)}, \end{aligned}$$

where the final inequality holds for any $\rho \in [0, 1]$, with the following optimal weights:

$$\sum_{j=4}^K w_{2j} = 1, \quad \text{and} \quad \sum_{j=4}^K w_{3j} = 0.$$

Notice that the smallest gap $\Delta = \psi(\{2, 3\}) - \psi(\{1, 2\})$ for the bandit instance governed by Σ simplifies to

$$\begin{aligned} \Delta &= \frac{1}{(1-\rho^4)} \left[(K-3)(\rho^2 + 3\rho^4 + 2\rho^6 - 2\rho^7) \right. \\ &\quad \left. + (2\rho^4 + 3\rho^6 - \rho^2) \right]. \end{aligned}$$

A simple calculation yields $\frac{\rho^4}{4(1+\rho^2)} \leq \Delta$, which implies $\mathbb{E}[\tau_\delta] \geq \frac{\log(\frac{1}{2.4\delta})}{\Delta}$. \square

VII. CONCLUSIONS

For the problem of estimation of the MSE of a given subset, with a multivariate Gaussian model, we proposed a natural estimator, and derived tail bounds that exponentially concentrate. Next, we framed the estimation problem with bandit feedback in the best-subset identification setting, and proposed a variant of the successive elimination technique. Finally, we also derived a minimax lower bound to understand the fundamental limit on the sample complexity of the aforementioned estimation problem with bandit feedback.

REFERENCES

- [1] C. Guestrin, A. Krause, and A. P. Singh, "Near-optimal sensor placements in gaussian processes," in *International conference on Machine learning*, 2005, pp. 265–272.
- [2] A. Krause, J. Leskovec, C. Guestrin, J. VanBriesen, and C. Faloutsos, "Efficient Sensor Placement Optimization for Securing Large Water Distribution Networks," *Journal of Water Resources Planning and Management*, vol. 134, no. 6, pp. 516–526, 2008.
- [3] U. Paul, L. Ortiz, S. R. Das, G. Fusco, , and M. M. Buddhikot, "Learning probabilistic models of cellular network traffic with applications to resource management," in *IEEE International Symposium on Dynamic Spectrum Access Networks*, 2014, pp. 82–91.
- [4] M. J. Wainwright, *High-dimensional statistics: A non-asymptotic viewpoint*. Cambridge University Press, 2019, vol. 48.
- [5] T. Lattimore and C. Szepesvári, *Bandit algorithms*. Cambridge University Press, 2020.
- [6] E. Even-Dar, S. Mannor, and Y. Mansour, "Pac bounds for multi-armed bandit and markov decision processes," in *International Conference on Computational Learning Theory*. Springer, 2002, pp. 255–270.
- [7] C. Y. Liu and S. Bubeck, "Most correlated arms identification," in *Conference on Learning Theory*, 2014, pp. 623–637.
- [8] V. P. Boda and L. A. Prashanth, "Correlated bandits or: How to minimize mean-squared error online," in *International Conference on Machine Learning*. PMLR, 2019, pp. 686–694.
- [9] S. Gupta, G. Joshi, and O. Yağan, "Correlated multi-armed bandits with a latent random source," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 3572–3576.
- [10] A. Erraqabi, A. Lazaric, M. Valko, E. Brunskill, and Y. Liu, "Trading off rewards and errors in multi-armed bandits," in *International Conference on Artificial Intelligence and Statistics*, vol. 54. PMLR, 2017, pp. 709–717.
- [11] D. Sen, L. Prashanth, and A. Gopalan, "Adaptive estimation of random vectors with bandit feedback," Indian Institute of Technology Madras, Tech. Rep., 2023. [Online]. Available: <http://www.cse.iitm.ac.in/~prashla/CorrelatedBanditsICC.pdf>
- [12] B. Hajek, "Notes for ECE 534: an exploration of random processes for engineers," *Univ. of Illinois at Urbana-Champaign*, 2009.
- [13] T. T. Cai, C. H. Zhang, and H. H. Zhou, "Optimal rates of convergence for covariance matrix estimation," *The Annals of Statistics*, vol. 38, no. 4, p. 2118–2144, Aug 2010.
- [14] E. Kaufmann, O. Cappé, and A. Garivier, "On the complexity of best arm identification in multi-armed bandit models," *The Journal of Machine Learning Research*, 2015.
- [15] P. Rigollet and J. Hütter, "High dimensional statistics," *Lecture notes for course 18S997*, 2019.

APPENDIX I
PROOFS

A. MSE Estimation: Non-adaptive setting

In this section, we prove the MSE concentration result for the non-adaptive setting. For ease of readability, we restate the concentration bound in Proposition 1 from the main paper.

Proposition 3. *Assume (A0) and (A1). Let $n_{AA}, n_{A'A}, n_{AA'}, n_{A'A'}$ denote the number of samples used to form $\widehat{\Sigma}_{AA}, \widehat{\Sigma}_{A'A}, \widehat{\Sigma}_{AA'},$ and $\Sigma_{A'A'}$, respectively. Let $\zeta = M_0 \min\left(\sqrt{\frac{m+\log(\frac{1}{\delta})}{n_{AA}}}, \frac{m+\log(\frac{1}{\delta})}{n_{AA}}\right)$. Then, for any $0 < \epsilon < \eta \triangleq \min(2K, \lambda_{\min}(\Sigma_{AA}))$, the MSE estimate $\widehat{\psi}(A)$ defined by (6) satisfies*

$$\mathbb{P}\left(|\widehat{\psi}(A) - \psi(A)| \geq \epsilon\right) \leq (13m(K-m) + \exp(m+K)) \exp\left(-\frac{n'}{m(K-m)^2(1+\eta)^3}\right) \min\left(\frac{\epsilon}{12\sqrt{2}G_0}, \frac{\epsilon^2}{G_0^2}\right) + (2m(K-m)) \exp\left(-\frac{n'}{72(m(K-m)^2)^2(1+\eta)^3 C_3^2}\right),$$

where $n' = \min(n_{AA}, n_{A'A}, n_{AA'}, n_{A'A'})$, $c = \frac{1}{[\lambda_{\min}(\Sigma_{AA}) - \epsilon]}$, $C_1 = \frac{108\sqrt{2}}{l} \left(c + \frac{1}{M_1}\right)$, $C_2 = \left(\frac{3cM_0}{M_1}\right)$, $C_3 = \left(c + \frac{1}{M_1}\right)$, and $G_0 = \max\left(\sqrt{m(K-m)^2(1+\eta)} C_1, \sqrt{m(K-m)^2(1+\eta)} C_2, 3(K-m)M_0\right)$.

Proof. Notice that

$$\widehat{\psi}(A) - \psi(A) = \text{Tr}\left(\left(\widehat{\Sigma}_{A'A'} - \Sigma_{A'A'}\right) - \left(\widehat{\Sigma}_{A'A}(\widehat{\Sigma}_{AA}^+)^{-1}\widehat{\Sigma}_{AA'} - \Sigma_{A'A}\Sigma_{AA}^{-1}\Sigma_{AA'}\right)\right). \quad (18)$$

The second term on the RHS of (18), can be re-written as follows:

$$\begin{aligned} (II) &= \widehat{\Sigma}_{A'A} \left(\widehat{\Sigma}_{AA}^+\right)^{-1} \widehat{\Sigma}_{AA'} - \Sigma_{A'A} (\Sigma_{AA})^{-1} \Sigma_{AA'} \\ &= \widehat{\Sigma}_{A'A} \left(\left(\widehat{\Sigma}_{AA}^+\right)^{-1} \widehat{\Sigma}_{AA'} - \Sigma_{AA}^{-1} \Sigma_{AA'}\right) + \left(\widehat{\Sigma}_{A'A} - \Sigma_{A'A}\right) \Sigma_{AA}^{-1} \Sigma_{AA'} \\ &= \widehat{\Sigma}_{A'A} \left(\left(\widehat{\Sigma}_{AA}^+\right)^{-1} \left(\widehat{\Sigma}_{AA'} - \Sigma_{AA'}\right) + \left(\left(\widehat{\Sigma}_{AA}^+\right)^{-1} - \Sigma_{AA}^{-1}\right) \Sigma_{AA'}\right) \\ &\quad + \left(\widehat{\Sigma}_{A'A} - \Sigma_{A'A}\right) \Sigma_{AA}^{-1} \Sigma_{AA'} \\ &= \widehat{\Sigma}_{A'A} \left(\widehat{\Sigma}_{AA}^+\right)^{-1} \left(\widehat{\Sigma}_{AA'} - \Sigma_{AA'}\right) + \widehat{\Sigma}_{A'A} \left(\left(\widehat{\Sigma}_{AA}^+\right)^{-1} - \Sigma_{AA}^{-1}\right) \Sigma_{AA'} \\ &\quad + \left(\widehat{\Sigma}_{A'A} - \Sigma_{A'A}\right) \Sigma_{AA}^{-1} \Sigma_{AA'}. \end{aligned} \quad (19)$$

Using the definition of the positive definite estimator $\widehat{\Sigma}^+$, we have

$$\|\widehat{\Sigma}_{AA}^+ - \Sigma_{AA}\|_2 \leq \|\widehat{\Sigma}_{AA}^+ - \widehat{\Sigma}_{AA}\| + \|\widehat{\Sigma}_{AA} - \Sigma_{AA}\|_2 \leq 2\zeta + \|\widehat{\Sigma}_{AA} - \Sigma_{AA}\|_2. \quad (20)$$

Using Theorem 5.7 of [15] in conjunction with (A1), w.p. $(1-\delta)$, we have

$$\begin{aligned} \|\widehat{\Sigma}_{AA} - \Sigma_{AA}\|_2 &\leq \|\Sigma_{AA}\|_2 \min\left(\sqrt{\frac{m+\log(\frac{1}{\delta})}{n_{AA}}}, \frac{m+\log(\frac{1}{\delta})}{n_{AA}}\right) \leq M_0 \min\left(\sqrt{\frac{m+\log(\frac{1}{\delta})}{n_{AA}}}, \frac{m+\log(\frac{1}{\delta})}{n_{AA}}\right), \text{ and} \\ \|\widehat{\Sigma}_{A'A'} - \Sigma_{A'A'}\|_2 &\leq M_0 \min\left(\sqrt{\frac{K-m+\log(\frac{1}{\delta})}{n_{A'A'}}}, \frac{K-m+\log(\frac{1}{\delta})}{n_{A'A'}}\right). \end{aligned} \quad (21)$$

With $c = \frac{1}{[\lambda_{\min}(\Sigma_{AA}) - \epsilon]}$, consider the event

$$\tilde{\mathcal{B}} = \{\sigma_i^2 - \epsilon \leq \hat{\sigma}_i^2 \leq \sigma_i^2 + \epsilon, i = [K], \rho_{i_k j} - \epsilon \leq \hat{\rho}_{i_k j} \leq \rho_{i_k j} + \epsilon, \text{ for } (k, j) \in [K], k \neq j, \|\left(\widehat{\Sigma}_{AA}^+\right)^{-1}\|_2 \leq c\}. \quad (22)$$

On the event $\tilde{\mathcal{B}}$, w.p. $(1-\delta)$, we have

$$\|\left(\widehat{\Sigma}_{AA}^+\right)^{-1} - \Sigma_{AA}^{-1}\|_2 = \|\left(\widehat{\Sigma}_{AA}\right)^{-1} (\Sigma_{AA} - \widehat{\Sigma}_{AA}^+) \Sigma_{AA}^{-1}\|_2 \leq \|\left(\widehat{\Sigma}_{AA}^+\right)^{-1}\|_2 \|\Sigma_{AA} - \widehat{\Sigma}_{AA}^+\|_2 \|\Sigma_{AA}^{-1}\|_2$$

$$\begin{aligned}
&\leq \left(\frac{c}{M_1}\right) \left(2\zeta + M_0 \min\left(\sqrt{\frac{m + \log(\frac{1}{\delta})}{n_{AA}}}, \frac{m + \log(\frac{1}{\delta})}{n_{AA}}\right)\right) \\
&= \frac{3cM_0}{M_1} \min\left(\sqrt{\frac{m + \log(\frac{1}{\delta})}{n_{AA}}}, \frac{m + \log(\frac{1}{\delta})}{n_{AA}}\right),
\end{aligned}$$

where we used (20), (21), and substituted the value of ζ specified in the proposition statement.

Letting $\eta = \min(2K, \lambda_{\min}(\Sigma_{AA}))$, we obtain $\|\widehat{\Sigma}_{AA'}\|_2 \leq \sqrt{m(K-m)}(1+\eta)$, (since $\hat{\sigma}_j^2 \leq \sigma_j^2 + \epsilon \leq 1 + \eta$ on $\tilde{\mathcal{B}}$) and $\|\Sigma_{AA'}\|_2 \sqrt{m(K-m)}(1+\eta)$. Similarly, $\|\widehat{\Sigma}_{A'A}\|_2 \leq \sqrt{m(K-m)}(1+\eta)$ and $\|\Sigma_{A'A}\|_2 \leq \sqrt{m(K-m)}(1+\eta)$. Thus,

$$\begin{aligned}
&\mathbb{P}\left(\|\widehat{\Sigma}_{AA'} - \Sigma_{AA'}\|_2^2 \geq \epsilon^2, \tilde{\mathcal{B}}\right) \leq \mathbb{P}\left(\|\widehat{\Sigma}_{AA'} - \Sigma_{AA'}\|_F^2 \geq \epsilon^2, \tilde{\mathcal{B}}\right) \\
&= \sum_{k=1}^m \sum_{j=m+1}^{K-m} \mathbb{P}\left(\left|\hat{\rho}_{jik} \hat{\sigma}_{i_k} \hat{\sigma}_j - \rho_{jik} \sigma_{i_k} \sigma_j\right| \geq \frac{\epsilon}{\sqrt{m(K-m)}}, \tilde{\mathcal{B}}\right) \\
&= \sum_{k=1}^m \sum_{j=m+1}^{K-m} \mathbb{P}\left(\left|\hat{\rho}_{jik} \hat{\sigma}_{i_k} (\hat{\sigma}_j - \sigma_j) + \hat{\rho}_{jik} \sigma_j (\hat{\sigma}_{i_k} - \sigma_{i_k}) + \sigma_{i_k} \sigma_j (\hat{\rho}_{jik} - \rho_{jik})\right| \geq \frac{\epsilon}{\sqrt{m(K-m)}}, \tilde{\mathcal{B}}\right) \\
&\leq \sum_{k=1}^m \sum_{j=m+1}^{K-m} \left(\mathbb{P}\left(\hat{\rho}_{jik} \hat{\sigma}_{i_k} (\hat{\sigma}_j - \sigma_j) \geq \frac{\epsilon}{3\sqrt{m(K-m)}}, \tilde{\mathcal{B}}\right)\right. \\
&\quad \left. + \mathbb{P}\left(\hat{\rho}_{jik} \sigma_j (\hat{\sigma}_{i_k} - \sigma_{i_k}) \geq \frac{\epsilon}{3\sqrt{m(K-m)}}, \tilde{\mathcal{B}}\right) + \mathbb{P}\left(\sigma_{i_k} \sigma_j (\hat{\rho}_{jik} - \rho_{jik}) \geq \frac{\epsilon}{3\sqrt{m(K-m)}}, \tilde{\mathcal{B}}\right)\right) \\
&\leq \sum_{k=1}^m \sum_{j=m+1}^{K-m} \left(\mathbb{P}\left((\hat{\sigma}_j - \sigma_j) \geq \frac{\epsilon}{3\sqrt{m(K-m)}(1+\eta)}\right) + \mathbb{P}\left((\hat{\sigma}_{i_k} - \sigma_{i_k}) \geq \frac{\epsilon}{3\sqrt{m(K-m)}(1+\eta)}\right)\right) \\
&\quad \left. + \mathbb{P}\left((\hat{\rho}_{jik} - \rho_{jik}) \geq \frac{\epsilon}{3\sqrt{m(K-m)}(1+\eta)}, \tilde{\mathcal{B}}\right)\right) \left(\text{since } \hat{\sigma}_j^2 \leq \sigma_j^2 + \epsilon \leq 1 + \eta, \text{ on } \tilde{\mathcal{B}}, \sigma_i^2, \hat{\rho}_{ij} \leq 1\right) \\
&\leq \sum_{k=1}^m \sum_{j=m+1}^{K-m} \left(\exp\left(-\frac{n_{AA'} \epsilon^2}{72m(K-m)(1+\eta)}\right) + \exp\left(-\frac{n_{AA'} \epsilon^2}{72m(K-m)(1+\eta)}\right)\right) \\
&\quad \left. + 13 \exp\left(-\frac{n_{AA'}}{8} \frac{1}{36(1+\eta)} \min\left(\frac{l\epsilon}{9\sqrt{m(K-m)}(1+\eta)}, \frac{l^2 \epsilon^2}{81m(K-m)(1+\eta)^2}\right)\right)\right) \\
&\leq (m(K-m)) \left(2 \exp\left(-\frac{n_{AA'} \epsilon^2}{72m(K-m)(1+\eta)}\right)\right. \\
&\quad \left. + 13 \exp\left(-\frac{n_{AA'}}{8} \frac{1}{36(1+\eta)} \min\left(\frac{l\epsilon}{9\sqrt{m(K-m)}(1+\eta)}, \frac{l^2 \epsilon^2}{81m(K-m)(1+\eta)^2}\right)\right)\right). \tag{23}
\end{aligned}$$

Now, using (23), w.p. $(1 - \delta)$, we have

$$\begin{aligned}
&\|\widehat{\Sigma}_{AA'} - \Sigma_{AA'}\|_2 \\
&\leq \left(\frac{1+\eta}{l}\right) \left(108\sqrt{2m(K-m)}\right) \min\left(\sqrt{\frac{(1+\eta) \log\left(\frac{13m(K-m)}{\delta}\right)}{n_{AA'}}}, \right. \\
&\quad \left.\left(\frac{12\sqrt{2}(1+\eta) \log\left(\frac{13m(K-m)}{\delta}\right)}{n_{AA'}}\right)\right) + \sqrt{\frac{72m(K-m)(1+\eta) \log\left(\frac{2m(K-m)}{\delta}\right)}{n_{AA'}}}.
\end{aligned}$$

Similarly, w.p. $(1 - \delta)$, we obtain

$$\|\widehat{\Sigma}_{A'A} - \Sigma_{A'A}\|_2$$

$$\leq \left(\frac{1+\eta}{l}\right) \left(108\sqrt{2m(K-m)}\right) \min \left(\sqrt{\frac{(1+\eta) \log\left(\frac{13m(K-m)}{\delta}\right)}{n_{A'A}}}, \right. \\ \left. \left(\frac{12\sqrt{2}(1+\eta) \log\left(\frac{13m(K-m)}{\delta}\right)}{n_{A'A}}\right) \right) + \sqrt{\frac{72m(K-m)(1+\eta) \log\left(\frac{2m(K-m)}{\delta}\right)}{n_{A'A}}}.$$

Now, the term (II) can be bounded on the event $\tilde{\mathcal{B}}$, w.p. $(1-\delta)$, as follows:

$$(II) \leq \|\widehat{\Sigma}_{A'A}\|_2 \|\widehat{\Sigma}_{AA}^+\|_2^{-1} \|\widehat{\Sigma}_{AA'} - \Sigma_{AA'}\|_2 + \|\widehat{\Sigma}_{A'A}\|_2 \|\widehat{\Sigma}_{AA'}\|_2 \|\widehat{\Sigma}_{AA}^+\|_2^{-1} - \Sigma_{AA}^{-1}\|_2 \\ + \|\Sigma_{AA'}\|_2 \|\Sigma_{AA}^{-1}\|_2 \|\widehat{\Sigma}_{A'A} - \Sigma_{A'A}\|_2 \\ \leq 108\sqrt{2}(cm(K-m)) \left(\frac{(1+\eta)^2}{l}\right) \min \left(\sqrt{\frac{(1+\eta) \log\left(\frac{13m(K-m)}{\delta}\right)}{n_{AA'}}}, \right. \\ \left. \left(\frac{12\sqrt{2}(1+\eta) \log\left(\frac{13m(K-m)}{\delta}\right)}{n_{AA'}}\right) \right) \\ + c\sqrt{m(K-m)}(1+\eta) \left(\sqrt{\frac{72m(K-m)(1+\eta) \log\left(\frac{2m(K-m)}{\delta}\right)}{n_{AA'}}} \right) \\ + 3(m(K-m)(1+\eta)^2) \left(\frac{c}{M_1}\right) \left(2\zeta + M_0 \min \left(\sqrt{\frac{m + \log(\frac{1}{\delta})}{n_{AA}}}, \frac{m + \log(\frac{1}{\delta})}{n_{AA}} \right) \right) \\ + 108\sqrt{2} \left(\frac{m(K-m)}{M_1}\right) \left(\frac{(1+\eta)^2}{l}\right) \min \left(\sqrt{\frac{(1+\eta) \log\left(\frac{13m(K-m)}{\delta}\right)}{n_{A'A}}}, \right. \\ \left. \left(\frac{12\sqrt{2}(1+\eta) \log\left(\frac{13m(K-m)}{\delta}\right)}{n_{A'A}}\right) \right) \\ + \left(\frac{\sqrt{m(K-m)}}{M_1}\right) (1+\eta) \left(\sqrt{\frac{72m(K-m)(1+\eta) \log\left(\frac{2m(K-m)}{\delta}\right)}{n_{A'A}}} \right) \\ \leq (m(K-m)(1+\eta)^2) \left(C_1 \min \left(\sqrt{\frac{(1+\eta) \log\left(\frac{13m(K-m)}{\delta}\right)}{n'}}, \right. \right. \\ \left. \left. \left(\frac{12\sqrt{2}(1+\eta) \log\left(\frac{13m(K-m)}{\delta}\right)}{n'}\right) \right) \right) + C_2 \min \left(\sqrt{\frac{m + \log(\frac{1}{\delta})}{n'}}, \frac{m + \log(\frac{1}{\delta})}{n'} \right) \\ + \sqrt{m(K-m)}(1+\eta) C_3 \left(\sqrt{\frac{72m(K-m)(1+\eta) \log\left(\frac{2m(K-m)}{\delta}\right)}{n'}} \right),$$

where $n' = \min(n_{AA}, n_{A'A}, n_{AA'}, n_{A'A})$, $C_1 = \frac{108\sqrt{2}}{l} \left(c + \frac{1}{M_1}\right)$, $C_2 = \left(\frac{3cM_0}{M_1}\right)$ and $C_3 = \left(c + \frac{1}{M_1}\right)$.

From the foregoing,

$$\mathbb{P} \left(|\widehat{\psi}(A) - \psi(A)| \geq \epsilon, \tilde{\mathcal{B}} \right)$$

$$\begin{aligned}
&\leq 13(m(K-m)) \exp\left(-\frac{n'}{m(K-m)^2(1+\eta)^3} \min\left(\frac{\epsilon}{12\sqrt{2}C_1}, \frac{\epsilon^2}{m(K-m)^2(1+\eta)^2C_1^2}\right)\right) \\
&\quad + \exp\left(m - \frac{n'}{m(K-m)^2(1+\eta)^2} \min\left(\frac{\epsilon}{C_2}, \frac{\epsilon^2}{m(K-m)^2(1+\eta)^2C_2^2}\right)\right) \\
&\quad + (2m(K-m)) \exp\left(-\frac{n'}{72(m(K-m)^2)^2(1+\eta)^3} \frac{\epsilon^2}{C_3^2}\right) \\
&\quad + \exp\left((K-m) - n' \min\left(\frac{\epsilon}{(K-m)M_0}, \frac{\epsilon^2}{(K-m)^2M_0^2}\right)\right) + \frac{6m(K-m)(1+\eta)^2c\zeta}{M_1}. \tag{24}
\end{aligned}$$

Let $\lambda_{\min}(\Sigma_{AA})$ and $\lambda_{\min}(\widehat{\Sigma}_{AA}^+)$ be the smallest eigenvalues of Σ_{AA} and $\widehat{\Sigma}_{AA}^+$ respectively. Then, for $0 < \epsilon < \eta$, we have

$$\begin{aligned}
\mathbb{P}(\lambda_{\min}(\widehat{\Sigma}_{AA}^+) \leq \lambda_{\min}(\Sigma_{AA}) - \epsilon) &= \mathbb{P}(\lambda_{\min}(\widehat{\Sigma}_{AA}^+) \leq 1/c) = \mathbb{P}(1/\lambda_{\min}(\widehat{\Sigma}_{AA}^+) \geq c) \\
&= \mathbb{P}(\|\widehat{\Sigma}_{AA}^+\|_2 \geq c).
\end{aligned}$$

Using a corollary of the Weyl's theorem (cf. p. 161 of [4]), we obtain

$$\mathbb{P}(\lambda_{\min}(\widehat{\Sigma}_{AA}^+) - \lambda_{\min}(\Sigma_{AA}) \geq \epsilon) = \mathbb{P}(\|\widehat{\Sigma}_{AA}^+\|_2 \geq c) \leq \mathbb{P}(\|\widehat{\Sigma}_{AA}^+ - \Sigma_{AA}\|_2 \geq \epsilon).$$

From (21) and (20), w.p. at least $(1 - \delta)$, we have

$$\|\widehat{\Sigma}_{AA}^+ - \Sigma_{AA}\|_2 \leq 2\zeta + M_0 \min\left(\sqrt{\frac{m + \log(\frac{1}{\delta})}{n_{AA}}}, \frac{m + \log(\frac{1}{\delta})}{n_{AA}}\right) = 3M_0 \min\left(\sqrt{\frac{m + \log(\frac{1}{\delta})}{n_{AA}}}, \frac{m + \log(\frac{1}{\delta})}{n_{AA}}\right),$$

where the final equality is obtained by substituting the value of ζ specified in the proposition statement. Hence,

$$\mathbb{P}(\tilde{\mathcal{B}}') \leq \exp\left(m - n_{AA} \min\left(\frac{\epsilon}{3M_0}, \frac{\epsilon^2}{9M_0^2}\right)\right). \tag{25}$$

Combining (24) and (25), we obtain

$$\begin{aligned}
\mathbb{P}\left(|\widehat{\psi}(A) - \psi(A)| \geq \epsilon\right) &\leq \mathbb{P}\left(|\widehat{\psi}(A) - \psi(A)| \geq \epsilon, \tilde{\mathcal{B}}\right) + \mathbb{P}(\tilde{\mathcal{B}}') \\
&\leq 13(m(K-m)) \exp\left(-\frac{n'}{m(K-m)^2(1+\eta)^3} \min\left(\frac{\epsilon}{12\sqrt{2}C_1}, \frac{\epsilon^2}{m(K-m)^2(1+\eta)^2C_1^2}\right)\right) \\
&\quad + \exp\left(m - \frac{n'}{m(K-m)^2(1+\eta)^2} \min\left(\frac{\epsilon}{C_2}, \frac{\epsilon^2}{m(K-m)^2(1+\eta)^2C_2^2}\right)\right) \\
&\quad + (2m(K-m)) \exp\left(-\frac{n'}{72(m(K-m)^2)^2(1+\eta)^3} \frac{\epsilon^2}{C_3^2}\right) \\
&\quad + \exp\left((K-m) - n' \min\left(\frac{\epsilon}{(K-m)M_0}, \frac{\epsilon^2}{(K-m)^2M_0^2}\right)\right) + \exp\left(m - n_{AA} \min\left(\frac{\epsilon}{3M_0}, \frac{\epsilon^2}{9M_0^2}\right)\right) \\
&\leq (13m(K-m) + \exp(m+K)) \exp\left(-\frac{n'}{m(K-m)^2(1+\eta)^3} \min\left(\frac{\epsilon}{12\sqrt{2}G_0}, \frac{\epsilon^2}{G_0^2}\right)\right) \\
&\quad + (2m(K-m)) \exp\left(-\frac{n'}{72(m(K-m)^2)^2(1+\eta)^3} \frac{\epsilon^2}{C_3^2}\right),
\end{aligned}$$

where $G_0 = \max\left(\sqrt{m(K-m)^2(1+\eta)} C_1, \sqrt{m(K-m)^2(1+\eta)} C_2, 3(K-m) M_0\right)$. \square

B. MSE Estimation: Adaptive setting

We restate the main result in the adaptive setting for the sake of readability.

Proposition 4. *Assume (A0) and (A1). Set the projection parameter ζ as follows:*

$$\zeta = \sqrt{\frac{(1+\eta)^3(m^2-m)}{n''l^2}} \sqrt{\log\left(\frac{15(m^2-m)}{\delta}\right)} + \sqrt{\frac{m \log\left(\frac{m}{\delta}\right)}{n''}}.$$

Then, for any $0 < \epsilon < \eta \triangleq \min(2K, \lambda_{\min}(\Sigma_{AA}))$, the MSE estimate formed using (10) satisfies

$$\mathbb{P}\left(|\widehat{\psi}(A) - \psi(A)| \geq \epsilon\right) \leq K \left((13m+1) \exp\left(-\frac{n''}{G_1} \min\left(\frac{l\epsilon}{12\sqrt{2}G_2}, \frac{l^2\epsilon^2}{G_2^2 m(1+\eta)^2}\right)\right) + (30m^2 - 26m) \exp\left(-\frac{n''l^2\epsilon^2}{G_3(m^4 - m^2)(1+\eta)^7}\right) \right),$$

where $n'' = \min(n_i, n_j, n_{ij}, (i, j) \in [K], i \neq j)$, C_4 is a universal constant, $C_5 = 108\sqrt{2}\left(c + \frac{1}{M_1}\right)$, $C_6 = \left(\frac{3c}{M_1}\right)$, $C_7 = \left(c + \frac{1}{M_1}\right)$, $G_1 = \max(8, m(1+\eta)^3)$, $G_2 = \max(1, C_5)$, and $G_3 = \max(C_4, C_4 C_6^2, 72 C_7^2)$.

The proof of the result above requires two lemmas, which we state and prove below.

Lemma 1. Under conditions of Proposition 4, we have

$$\mathbb{P}\left(\|\widehat{\Sigma}_{AA} - \Sigma_{AA}\|_2 \geq \epsilon, \tilde{\mathcal{B}}\right) \leq m \exp\left(-\frac{n''\epsilon^2}{C_4 m}\right) + 15(m^2 - m) \exp\left(-\frac{n''l^2\epsilon^2}{C_4(1+\eta)^3(m^2 - m)}\right),$$

where the symbols are as defined in the statement of Proposition 2.

Proof. Notice that

$$\begin{aligned} \|\widehat{\Sigma}_{AA} - \Sigma_{AA}\|_F^2 &= \sum_{(i,j) \in A} \left(\widehat{\Sigma}_{AA}(i,j) - \Sigma_{AA}(i,j)\right)^2 \\ &= \sum_{i \neq j, (i,j) \in A} \left(\widehat{\Sigma}_{AA}(i,j) - \Sigma_{AA}(i,j)\right)^2 + \sum_{i=j, (i,j) \in A} \left(\widehat{\Sigma}_{AA}(i,j) - \Sigma_{AA}(i,j)\right)^2 \\ &= \sum_{i \neq j} (\hat{\rho}_{ij}\hat{\sigma}_i\hat{\sigma}_j - \rho_{ij}\sigma_i\sigma_j)^2 + \sum_{i=j} (\hat{\sigma}_i^2 - \sigma_i^2)^2. \end{aligned}$$

Now,

$$\begin{aligned} \mathbb{P}\left(\|\widehat{\Sigma}_{AA} - \Sigma_{AA}\|_2^2 \geq \epsilon^2, \tilde{\mathcal{B}}\right) &\leq \mathbb{P}\left(\|\widehat{\Sigma}_{AA} - \Sigma_{AA}\|_F^2 \geq \epsilon^2, \tilde{\mathcal{B}}\right) \\ &= \mathbb{P}\left(\left(\sum_{i \neq j} (\hat{\rho}_{ij}\hat{\sigma}_i\hat{\sigma}_j - \rho_{ij}\sigma_i\sigma_j)^2 + \sum_{i=j} (\hat{\sigma}_i^2 - \sigma_i^2)^2\right) \geq \epsilon^2, \tilde{\mathcal{B}}\right) \\ &\leq \mathbb{P}\left(\left(\sum_{i \neq j} (\hat{\rho}_{ij}\hat{\sigma}_i\hat{\sigma}_j - \rho_{ij}\sigma_i\sigma_j)^2\right) \geq \frac{\epsilon^2}{2}, \tilde{\mathcal{B}}\right) + \mathbb{P}\left(\left(\sum_{i=j} (\hat{\sigma}_i^2 - \sigma_i^2)^2\right) \geq \frac{\epsilon^2}{2}\right) \\ &\leq \sum_{i \neq j} \mathbb{P}\left((\hat{\rho}_{ij}\hat{\sigma}_i\hat{\sigma}_j - \rho_{ij}\sigma_i\sigma_j)^2 \geq \frac{\epsilon^2}{2(m^2 - m)}, \tilde{\mathcal{B}}\right) + \sum_{i=j} \mathbb{P}\left((\hat{\sigma}_i^2 - \sigma_i^2)^2 \geq \frac{\epsilon^2}{2m}\right) \\ &\leq \sum_{i \neq j} \mathbb{P}\left((\hat{\rho}_{ij}\hat{\sigma}_i(\hat{\sigma}_j - \sigma_j) + \hat{\rho}_{ij}\sigma_j(\hat{\sigma}_i - \sigma_i) + \sigma_i\sigma_j(\hat{\rho}_{ij} - \rho_{ij})) \geq \frac{\epsilon}{\sqrt{2(m^2 - m)}}, \tilde{\mathcal{B}}\right) \\ &\quad + \sum_{i=j} \mathbb{P}\left((\hat{\sigma}_i - \sigma_i) \geq \frac{\epsilon}{\sqrt{2m}}\right) \\ &\leq \sum_{i \neq j} \left(\mathbb{P}\left((\hat{\rho}_{ij}\hat{\sigma}_i(\hat{\sigma}_j - \sigma_j)) \geq \frac{\epsilon}{3\sqrt{2(m^2 - m)}}, \tilde{\mathcal{B}}\right) + \mathbb{P}\left((\hat{\rho}_{ij}\sigma_j(\hat{\sigma}_i - \sigma_i)) \geq \frac{\epsilon}{3\sqrt{2(m^2 - m)}}, \tilde{\mathcal{B}}\right) + \mathbb{P}\left(\sigma_i\sigma_j(\hat{\rho}_{ij} - \rho_{ij}) \geq \frac{\epsilon}{3\sqrt{2(m^2 - m)}}, \tilde{\mathcal{B}}\right)\right) \\ &\quad + \sum_{i=j} \mathbb{P}\left((\hat{\sigma}_i - \sigma_i) \geq \frac{\epsilon}{\sqrt{2m}}\right) \\ &\leq \sum_{i \neq j} \left(\mathbb{P}\left((\hat{\sigma}_j - \sigma_j) \geq \frac{\epsilon}{3\sqrt{2(1+\eta)(m^2 - m)}}, \tilde{\mathcal{B}}\right) + \mathbb{P}\left((\hat{\sigma}_i - \sigma_i) \geq \frac{\epsilon}{3\sqrt{2(1+\eta)(m^2 - m)}}, \tilde{\mathcal{B}}\right)\right) \end{aligned}$$

$$\begin{aligned}
& + \mathbb{P} \left((\hat{\rho}_{ij} - \rho_{ij}) \geq \frac{\epsilon}{3(1+\eta)\sqrt{2(m^2-m)}} \right) + \sum_{i=j} \exp \left(-\frac{n_i}{8} \frac{\epsilon^2}{2m} \right) \\
& (\text{since } \hat{\sigma}_i^2 \leq \sigma_i^2 + \epsilon \leq 1 + \eta \text{ on } \tilde{\mathcal{B}} \text{ where } \eta = \min(2K, \lambda_{\min}(\Sigma_{AA})), \hat{\rho}_{ij} \leq 1). \\
& \leq \sum_{i \neq j} \left(2 \exp \left(-\frac{\min(n_i, n_j)}{8} \frac{\epsilon^2}{18(1+\eta)(m^2-m)} \right) \right. \\
& \quad \left. + 13 \exp \left(-\frac{n_{ij}}{8} \frac{1}{36(1+\eta)} \min \left(\frac{l\epsilon}{9(1+\eta)\sqrt{2(m^2-m)}}, \frac{l^2\epsilon^2}{162(1+\eta)^2(m^2-m)} \right) \right) \right) \\
& \quad + \sum_{i=j} \exp \left(-\frac{n_i}{8} \frac{\epsilon^2}{2m} \right) \\
& \leq 2(m^2-m) \cdot \exp \left(-\frac{n''\epsilon^2}{C_4(1+\eta)(m^2-m)} \right) + m \cdot \exp \left(-\frac{n''}{C_4m} \right) \\
& \quad + 13(m^2-m) \exp \left(-\frac{n''l^2\epsilon^2}{C_4(1+\eta)^3(m^2-m)} \right) \\
& \leq 15(m^2-m) \exp \left(-\frac{n''l^2\epsilon^2}{C_4(1+\eta)^3(m^2-m)} \right) + m \exp \left(-\frac{n''}{C_4m} \right),
\end{aligned}$$

where $n'' = \min(n_i, n_j, n_{ij}, (i, j) \in [K], i \neq j)$ and C_4 is a universal constant.

Hence,

$$\mathbb{P} \left(\|\widehat{\Sigma}_{AA} - \Sigma_{AA}\|_2 \geq \epsilon, \tilde{\mathcal{B}} \right) \leq 15(m^2-m) \exp \left(-\frac{n''l^2\epsilon^2}{C_4(1+\eta)^3(m^2-m)} \right) + m \exp \left(-\frac{n''}{C_4m} \right). \quad (26)$$

□

Lemma 2. Under conditions of Proposition 4, we have

$$\begin{aligned}
\mathbb{P} \left(\left\| \left(\widehat{\Sigma}_{AA}^+ \right)^{-1} - \Sigma_{AA}^{-1} \right\|_2 \geq \epsilon, \tilde{\mathcal{B}} \right) & \leq 15(m^2-m) \exp \left(-\frac{n''M_1^2l^2\epsilon^2}{9C_4(1+\eta)^3(m^2-m)c^2} \right) \\
& \quad + m \exp \left(-\frac{n''M_1^2\epsilon^2}{9C_4m c^2} \right),
\end{aligned}$$

where $\tilde{\mathcal{B}}$, as defined in (22), is an event that satisfies

$$\mathbb{P}(\tilde{\mathcal{B}}^c) \leq 15(m^2-m) \exp \left(-\frac{n''l^2\epsilon^2}{9C_4(1+\eta)^3(m^2-m)} \right) + m \exp \left(-\frac{n''}{9C_4m} \right). \quad (27)$$

In the above, the symbols are as defined in the statement of Proposition 4.

Proof. Using (26), we obtain the following bound, which holds w.p. $(1-\delta)$,

$$\|\widehat{\Sigma}_{AA} - \Sigma_{AA}\|_2 \leq \sqrt{\frac{C_4(1+\eta)^3(m^2-m)}{n''l^2}} \sqrt{\log \left(\frac{15(m^2-m)}{\delta} \right)} + \sqrt{\frac{C_4m \log \left(\frac{m}{\delta} \right)}{n''}}.$$

On the event $\tilde{\mathcal{B}}$, w.p. $(1-\delta)$, we have

$$\begin{aligned}
\left\| \left(\widehat{\Sigma}_{AA}^+ \right)^{-1} - \Sigma_{AA}^{-1} \right\|_2 & \leq \left(\frac{3c}{M_1} \right) \left(\sqrt{\frac{C_4(1+\eta)^3(m^2-m)}{n''l^2}} \sqrt{\log \left(\frac{15(m^2-m)}{\delta} \right)} \right. \\
& \quad \left. + \sqrt{\frac{C_4m \log \left(\frac{m}{\delta} \right)}{n''}} \right). \quad (28)
\end{aligned}$$

The inequality above follows by using (20) and

$$\begin{aligned}
\left\| \left(\widehat{\Sigma}_{AA}^+ \right)^{-1} - \Sigma_{AA}^{-1} \right\|_2 & = \left\| \left(\widehat{\Sigma}_{AA} \right)^{-1} (\Sigma_{AA} - \widehat{\Sigma}_{AA}^+) \Sigma_{AA}^{-1} \right\|_2 \\
& \leq \left\| \left(\widehat{\Sigma}_{AA}^+ \right)^{-1} \right\|_2 \|\Sigma_{AA} - \widehat{\Sigma}_{AA}^+\|_2 \|\Sigma_{AA}^{-1}\|_2.
\end{aligned}$$

From (28), we obtain

$$\mathbb{P}\left(\left\|\left(\widehat{\Sigma}_{AA}^+\right)^{-1}-\Sigma_{AA}^{-1}\right\|_2 \geq \epsilon, \tilde{\mathcal{B}}\right) \leq 15(m^2-m) \exp\left(-\frac{n'' M_1^2 l^2 \epsilon^2}{9 C_4 (1+\eta)^3 (m^2-m) c^2}\right) + m \exp\left(-\frac{n'' M_1^2 \epsilon^2}{9 C_4 m c^2}\right).$$

Using an argument similar to the one employed in the proof of Proposition 1 in conjunction with (26), we obtain

$$\mathbb{P}(\tilde{\mathcal{B}}') \leq 15(m^2-m) \exp\left(-\frac{n'' l^2 \epsilon^2}{9 C_4 (1+\eta)^3 (m^2-m)}\right) + m \exp\left(-\frac{n'' \epsilon^2}{9 C_4 m}\right). \quad (29)$$

□

Proof of Proposition 4:

Proof. From (9) and (10), we have

$$\widehat{\psi}(A) - \psi(A) = \sum_{j=1}^K \left[(\hat{\sigma}_j^2 - \sigma_j^2) - \left(\widehat{C}_j (\widehat{\Sigma}_{AA}^+)^{-1} \widehat{C}_j^\top - C_j (\Sigma_{AA}^{-1}) C_j^\top \right) \right]. \quad (30)$$

The second term on the RHS of (30) can be re-written as follows:

$$\begin{aligned} (II) &= \widehat{C}_j (\widehat{\Sigma}_{AA}^+)^{-1} \widehat{C}_j^\top - C_j (\Sigma_{AA}^{-1}) C_j^\top \\ &= \widehat{C}_j \left((\widehat{\Sigma}_{AA}^+)^{-1} \widehat{C}_j^\top - \Sigma_{AA}^{-1} C_j^\top \right) + (\widehat{C}_j - C_j) \Sigma_{AA}^{-1} C_j^\top \\ &= \widehat{C}_j \left((\widehat{\Sigma}_{AA}^+)^{-1} (\widehat{C}_j^\top - C_j^\top) + C_j^\top \left((\widehat{\Sigma}_{AA}^+)^{-1} - \Sigma_{AA}^{-1} \right) \right) + (\widehat{C}_j - C_j) \Sigma_{AA}^{-1} C_j^\top \\ &= \widehat{C}_j (\widehat{\Sigma}_{AA}^+)^{-1} (\widehat{C}_j^\top - C_j^\top) + \widehat{C}_j C_j^\top \left((\widehat{\Sigma}_{AA}^+)^{-1} - \Sigma_{AA}^{-1} \right) + (\widehat{C}_j - C_j) \Sigma_{AA}^{-1} C_j^\top. \end{aligned} \quad (31)$$

Letting $\eta = \min(2K, \lambda_{\min}(\Sigma_{AA}))$, we obtain $\|\widehat{C}_j\|_2 \leq \sqrt{m(1+\eta)^2} \leq \sqrt{m}(1+\eta)$. (since $\hat{\sigma}_j^2 \leq \sigma_j^2 + \epsilon \leq 1 + \eta$ on $\tilde{\mathcal{B}}, \hat{\rho}_{ji} \leq 1$) and $\|C_j\|_2 \leq \sqrt{m(1+\eta)^2} \leq \sqrt{m}(1+\eta)$. Similarly, $\|\widehat{C}_j^\top\|_2 \leq \sqrt{m}(1+\eta)$ and $\|C_j^\top\|_2 \leq \sqrt{m}(1+\eta)$.

$$\begin{aligned} &\mathbb{P}\left(\|\widehat{C}_j - C_j\|_2^2 \geq \epsilon^2, \tilde{\mathcal{B}}\right) \\ &\leq \mathbb{P}\left(\|\widehat{C}_j - C_j\|_F^2 \geq \epsilon^2, \tilde{\mathcal{B}}\right) = \sum_{k=1}^m \mathbb{P}\left(|\hat{\rho}_{ji_k} \hat{\sigma}_{i_k} \hat{\sigma}_j - \rho_{ji_k} \sigma_{i_k} \sigma_j| \geq \frac{\epsilon}{\sqrt{m}}, \tilde{\mathcal{B}}\right) \\ &= \sum_{k=1}^m \mathbb{P}\left(|\hat{\rho}_{ji_k} \hat{\sigma}_{i_k} (\hat{\sigma}_j - \sigma_j) + \hat{\rho}_{ji_k} \sigma_j (\hat{\sigma}_{i_k} - \sigma_{i_k}) + \sigma_{i_k} \sigma_j (\hat{\rho}_{ji_k} - \rho_{ji_k})| \geq \frac{\epsilon}{\sqrt{m}}, \tilde{\mathcal{B}}\right) \\ &\leq \sum_{k=1}^m \left(\mathbb{P}\left(\hat{\rho}_{ji_k} \hat{\sigma}_{i_k} (\hat{\sigma}_j - \sigma_j) \geq \frac{\epsilon}{3\sqrt{m}}, \tilde{\mathcal{B}}\right) + \mathbb{P}\left(\hat{\rho}_{ji_k} \sigma_j (\hat{\sigma}_{i_k} - \sigma_{i_k}) \geq \frac{\epsilon}{3\sqrt{m}}, \tilde{\mathcal{B}}\right) \right. \\ &\quad \left. + \mathbb{P}\left(\sigma_{i_k} \sigma_j (\hat{\rho}_{ji_k} - \rho_{ji_k}) \geq \frac{\epsilon}{3\sqrt{m}}, \tilde{\mathcal{B}}\right) \right) \\ &\leq \sum_{k=1}^m \left(\mathbb{P}\left((\hat{\sigma}_j - \sigma_j) \geq \frac{\epsilon}{3\sqrt{m(1+\eta)}}\right) + \mathbb{P}\left((\hat{\sigma}_{i_k} - \sigma_{i_k}) \geq \frac{\epsilon}{3\sqrt{m(1+\eta)}}\right) \right. \\ &\quad \left. + \mathbb{P}\left((\hat{\rho}_{ji_k} - \rho_{ji_k}) \geq \frac{\epsilon}{3(1+\eta)\sqrt{m}}, \tilde{\mathcal{B}}\right) \right) \left(\text{since } \hat{\sigma}_j^2 \leq \sigma_j^2 + \epsilon \leq 1 + \eta, \text{ on } \tilde{\mathcal{B}}, \sigma_i^2, \hat{\rho}_{ij} \leq 1 \right) \\ &\leq \sum_{k=1}^m \left(\exp\left(-\frac{n_j \epsilon^2}{72m(1+\eta)}\right) + \exp\left(-\frac{n_{i_k} \epsilon^2}{72m(1+\eta)}\right) \right. \\ &\quad \left. + 13 \exp\left(-\frac{n_{ji_k}}{8} \frac{1}{324\sqrt{m(1+\eta)^2}} \min\left(\frac{l\epsilon}{1}, \frac{l^2 \epsilon^2}{9\sqrt{m(1+\eta)}}\right)\right) \right) \\ &\leq m \left(\exp\left(-\frac{n'' \epsilon^2}{72m(1+\eta)}\right) + \exp\left(-\frac{n'' \epsilon^2}{72m(1+\eta)}\right) \right) \end{aligned}$$

$$+13 \exp \left(-\frac{n''}{8} \frac{1}{324\sqrt{m}(1+\eta)^2} \min \left(l\epsilon, \frac{l^2\epsilon^2}{9\sqrt{m}(1+\eta)} \right) \right), \quad (32)$$

where $n'' = \min(n_i, n_j, n_{ij}, (i, j) \in [K], i \neq j)$. Now, using (32), we obtain the following bound, which holds w.p. $(1 - \delta)$:

$$\begin{aligned} \|\widehat{C}_j - C_j\|_2 &\leq \left[\frac{108\sqrt{2m}(1+\eta)}{l} \right] \min \left(\sqrt{\frac{(1+\eta) \log \left(\frac{13m}{\delta} \right)}{n''}}, \left[\frac{12\sqrt{2}(1+\eta) \log \left(\frac{13m}{\delta} \right)}{n''} \right] \right) \\ &\quad + \sqrt{\frac{72m(1+\eta) \log \left(\frac{2m}{\delta} \right)}{n''}}. \end{aligned}$$

Similarly, w.p. $(1 - \delta)$,

$$\begin{aligned} \|\widehat{C}_j^\top - C_j^\top\|_2 &\leq \left[\frac{108\sqrt{2m}(1+\eta)}{l} \right] \min \left(\sqrt{\frac{(1+\eta) \log \left(\frac{13m}{\delta} \right)}{n''}}, \left[\frac{12\sqrt{2}(1+\eta) \log \left(\frac{13m}{\delta} \right)}{n''} \right] \right) \\ &\quad + \sqrt{\frac{72m(1+\eta) \log \left(\frac{2m}{\delta} \right)}{n''}}. \end{aligned}$$

On the event $\tilde{\mathcal{B}}$, using (28), we obtain the following bound, which holds w.p. $(1 - \delta)$:

$$\begin{aligned} \left\| \left(\widehat{\Sigma}_{AA}^+ \right)^{-1} - \Sigma_{AA}^{-1} \right\|_2 &\leq \left(\frac{3c}{M_1} \right) \left(\sqrt{\frac{C_4(1+\eta)^3(m^2-m)}{n''l^2}} \sqrt{\log \left(\frac{15(m^2-m)}{\delta} \right)} \right. \\ &\quad \left. + \sqrt{\frac{C_4m \log \left(\frac{m}{\delta} \right)}{n''}} \right). \end{aligned}$$

Now, the term (II) can be bounded on the event $\tilde{\mathcal{B}}$, w.p. $(1 - \delta)$, as follows:

$$\begin{aligned} (II) &\leq \|\widehat{C}_j\|_2 \left\| \left(\widehat{\Sigma}_{AA}^+ \right)^{-1} \right\|_2 \|\widehat{C}_j^\top - C_j^\top\|_2 + \|\widehat{C}_j\|_2 \|C_j^\top\|_2 \left\| \left(\widehat{\Sigma}_{AA}^+ \right)^{-1} - \Sigma_{AA}^{-1} \right\|_2 \\ &\quad + \|\widehat{C}_j - C_j\|_2 \|\Sigma_{AA}^{-1}\|_2 \|C_j^\top\|_2 \\ &\leq (108\sqrt{2}c) \left(\frac{m(1+\eta)^2}{l} \right) \left(\min \left(\sqrt{\frac{(1+\eta) \log \left(\frac{13m}{\delta} \right)}{n''}}, \left(\frac{12\sqrt{2}(1+\eta) \log \left(\frac{13m}{\delta} \right)}{n''} \right) \right) \right) \\ &\quad + (cm(1+\eta)) \left(\sqrt{\frac{72(1+\eta) \log \left(\frac{2m}{\delta} \right)}{n''}} \right) \\ &\quad + (m(1+\eta)^2) \left(\frac{3c}{M_1} \right) \left(\sqrt{\frac{C_4(1+\eta)^3(m^2-m)}{n''l^2}} \sqrt{\log \left(\frac{15(m^2-m)}{\delta} \right)} + \sqrt{\frac{C_4m \log \left(\frac{m}{\delta} \right)}{n''}} \right) \\ &\quad + (108\sqrt{2}M_1) \left(\frac{m(1+\eta)^2}{l} \right) \left(\min \left(\sqrt{\frac{(1+\eta) \log \left(\frac{13m}{\delta} \right)}{n''}}, \left(\frac{12\sqrt{2}(1+\eta) \log \left(\frac{13m}{\delta} \right)}{n''} \right) \right) \right) \\ &\quad + \frac{m(1+\eta)}{M_1} \left(\sqrt{\frac{72(1+\eta) \log \left(\frac{2m}{\delta} \right)}{n''}} \right) \\ &\leq C_5 \left(\frac{m(1+\eta)^2}{l} \right) \left(\min \left(\sqrt{\frac{(1+\eta) \log \left(\frac{13m}{\delta} \right)}{n''}}, \left(\frac{12\sqrt{2}(1+\eta) \log \left(\frac{13m}{\delta} \right)}{n''} \right) \right) \right) \\ &\quad + C_6 (m(1+\eta)^2) \left(\sqrt{\frac{C_4(1+\eta)^3(m^2-m)}{n''l^2}} \sqrt{\log \left(\frac{15(m^2-m)}{\delta} \right)} + \sqrt{\frac{C_4m \log \left(\frac{m}{\delta} \right)}{n''}} \right) \\ &\quad + C_7 (m(1+\eta)) \left(\sqrt{\frac{72(1+\eta) \log \left(\frac{2m}{\delta} \right)}{n''}} \right), \end{aligned}$$

where $C_5 = 108\sqrt{2} \left(c + \frac{1}{M_1}\right)$, $C_6 = \left(\frac{3c}{M_1}\right)$ and $C_7 = \left(c + \frac{1}{M_1}\right)$.

From the foregoing,

$$\begin{aligned}
& \mathbb{P} \left(|\widehat{\psi}(A) - \psi(A)| \geq \epsilon, \tilde{\mathcal{B}} \right) \\
& \leq K \left(\exp \left(-\frac{n''}{8} \min(\epsilon, \epsilon^2) \right) + 13m \exp \left(-\frac{n''}{m(1+\eta)^3} \min \left(\frac{l\epsilon}{12\sqrt{2} C_5}, \frac{l^2\epsilon^2}{C_5^2 m(1+\eta)^2} \right) \right) \right. \\
& \quad + 15(m^2 - m) \exp \left(-\frac{n'' l^2 \epsilon^2}{C_4 C_6^2 (m^4 - m^2)(1+\eta)^7} \right) + m \exp \left(-\frac{n'' \epsilon^2}{C_4 C_6^2 m^2 (1+\eta)^4} \right) \\
& \quad \left. + 2m \exp \left(-\frac{n''}{72 m^2 (1+\eta)^2 C_7^2} \right) \right). \tag{33}
\end{aligned}$$

From (29), we have

$$\mathbb{P}(\tilde{\mathcal{B}}') \leq 15(m^2 - m) \exp \left(-\frac{n'' l^2 \epsilon^2}{9 C_4 (1+\eta)^3 (m^2 - m)} \right) + m \exp \left(-\frac{n'' \epsilon^2}{9 C_4 m} \right).$$

Combining (33) and (29), we obtain,

$$\begin{aligned}
& \mathbb{P} \left(|\widehat{\psi}(A) - \psi(A)| \geq \epsilon \right) \\
& \leq K \left(\exp \left(-\frac{n''}{8} \min(\epsilon, \epsilon^2) \right) + 13m \exp \left(-\frac{n''}{m(1+\eta)^3} \min \left(\frac{l\epsilon}{12\sqrt{2} C_5}, \frac{l^2\epsilon^2}{C_5^2 m(1+\eta)^2} \right) \right) \right. \\
& \quad + 15(m^2 - m) \exp \left(-\frac{n'' l^2 \epsilon^2}{C_4 C_6^2 (m^4 - m^2)(1+\eta)^7} \right) + m \exp \left(-\frac{n'' \epsilon^2}{C_4 C_6^2 m^2 (1+\eta)^4} \right) \\
& \quad + 2m \exp \left(-\frac{n''}{72 m^2 (1+\eta)^2 C_7^2} \right) + 15(m^2 - m) \exp \left(-\frac{n'' l^2 \epsilon^2}{9 C_4 (1+\eta)^3 (m^2 - m)} \right) \\
& \quad \left. + m \exp \left(-\frac{n'' \epsilon^2}{9 C_4 m} \right) \right) \\
& \leq K \left((13m + 1) \exp \left(-\frac{n''}{G_1} \min \left(\frac{l\epsilon}{12\sqrt{2} G_2}, \frac{l^2\epsilon^2}{G_2^2 m(1+\eta)^2} \right) \right) \right. \\
& \quad \left. + (30m^2 - 26m) \exp \left(-\frac{n'' l^2 \epsilon^2}{G_3 (m^4 - m^2)(1+\eta)^7} \right) \right),
\end{aligned}$$

where $G_1 = \max(8, m(1+\eta)^3)$, $G_2 = \max(1, C_5)$ and $G_3 = \max(C_4, C_4 C_6^2, 72 C_7^2)$. \square

C. Successive elimination

For deriving the confidence width α_t used in the successive elimination algorithm for correlated bandits (see Section V), we start by deriving an alternative form of the bound on the MSE estimate stated in Proposition 2.

$$\begin{aligned}
& \mathbb{P} \left(|\widehat{\psi}(A) - \psi(A)| \geq \epsilon \right) \leq K \left((13m + 1) \exp \left(-\frac{n''}{G_1} \min \left(\frac{l\epsilon}{12\sqrt{2} G_2}, \frac{l^2\epsilon^2}{G_2^2 m(1+\eta)^2} \right) \right) \right. \\
& \quad \left. + (30m^2 - 26m) \exp \left(-\frac{n'' l^2 \epsilon^2}{G_3 (m^4 - m^2)(1+\eta)^7} \right) \right) \\
& \leq K \left((13m + 1) \exp \left(-\frac{n'' l^2 \epsilon^2}{12\sqrt{2} l G_1 G_2 \epsilon + G_1 G_2^2 m(1+\eta)^2} \right) \right. \\
& \quad \left. + (30m^2 - 26m) \exp \left(-\frac{n'' l^2 \epsilon^2}{G_3 (m^4 - m^2)(1+\eta)^7} \right) \right) \\
& \leq K \left((13m + 1) \exp \left(-\frac{n'' \epsilon^2}{c_1 + c_2 \epsilon} \right) + (30m^2 - 26m) \exp(-n'' c_3 \epsilon^2) \right) \\
& \leq K \left((13m + 1 + (30m^2 - 26m)) \exp \left(-\frac{n'' \min(c_3, 1) \epsilon^2}{c_1 + c_2 \epsilon} \right) \right),
\end{aligned}$$

where $c_1 = \frac{G_1 G_2^2 m(1+\eta)^2}{l^2}$, $c_2 = \frac{12\sqrt{2} G_1 G_2}{l}$ and $c_3 = \frac{l^2}{G_2 (m^4 - m^2) (1+\eta)^7}$.

Therefore,

$$\mathbb{P}\left(|\widehat{\psi}(A) - \psi(A)| \geq \epsilon\right) \leq K \left((13m + 1 + (30m^2 - 26m)) \exp\left(-\frac{n'' c_3 \epsilon^2}{c_1 + c_2 \epsilon}\right) \right), \text{ since } c_3 \leq 1. \quad (34)$$

From (34), w.p. $1 - \delta$, we obtain

$$|\widehat{\psi}(A) - \psi(A)| \leq \left(\frac{c_2 \log\left(\frac{K(13m+1+(30m^2-26m))}{\delta}\right)}{n'' c_3} + \sqrt{\frac{c_1 \log\left(\frac{K(13m+1+(30m^2-26m))}{\delta}\right)}{n'' c_3}} \right). \quad (35)$$

Now, from (35), we obtain the following form for the confidence width α_t :

$$\alpha_t = \left(\frac{c_2 \log\left(\frac{70 \binom{K}{m} K m^2 t^2}{\delta}\right)}{2c_3 t} + \sqrt{\frac{c_1 \log\left(\frac{70 \binom{K}{m} K m^2 t^2}{\delta}\right)}{2c_3 t}} \right). \quad (36)$$

For the sake of readability, we restate below the main result concerning the successive elimination algorithm in the correlated bandit framework.

Theorem 3. *The successive elimination algorithm is $(0, \delta)$ -PAC and w.p. at least $1 - \delta$, it's sample complexity is bounded by*

$$\mathcal{O}\left(\sum_{A \in \mathcal{A}} \frac{1}{\Delta(A)} \log\left(\frac{\binom{K}{m} K m^2 \log(\Delta(A)^{-1})}{\delta}\right)\right).$$

Proof.

$$\text{Define the event } E = \left\{ |\widehat{\psi}_{A_t} - \psi_A| < \alpha_t, \forall t = 1, 2, \dots \text{ and } \forall A_t \in \mathcal{A} \right\}.$$

We establish below that $\mathbb{P}(E') \leq \delta$.

$$\begin{aligned} \mathbb{P}(E') &= \mathbb{P}\left(\sum_{t=1}^{\infty} \sum_{\mathcal{A}} \left(|\widehat{\psi}_{A_t} - \psi_A| \geq \alpha_t\right)\right) \\ &\leq \sum_{t=1}^{\infty} \sum_{\mathcal{A}} \mathbb{P}\left(\left(|\widehat{\psi}_{A_t} - \psi_A| \geq \alpha_t\right)\right) \\ &\leq \sum_{t=1}^{\infty} \sum_{\mathcal{A}} 2K \left((13m + 1 + (30m^2 - 26m)) \exp\left(-\frac{n'' c_3 \alpha_t^2}{c_1 + c_2 \alpha_t}\right) \right) \\ &\leq \sum_{t=1}^{\infty} \sum_{\mathcal{A}} 2K \left((13m + 1 + (30m^2 - 26m)) \exp\left(-\frac{n'' \log\left(\frac{70 \binom{K}{m} (K m^2) t^2}{\delta}\right)}{2t}\right) \right) \\ &\leq \sum_{t=1}^{\infty} \sum_{\mathcal{A}} K \left((13m + 1 + (30m^2 - 26m)) \exp\left(-\log\left(\frac{70 \binom{K}{m} (K m^2) t^2}{\delta}\right)\right) \right) \\ &\quad \text{since } n'' \geq t \\ &\leq \sum_{t=1}^{\infty} \sum_{\mathcal{A}} K \left((13m + 1 + (30m^2 - 26m)) \right) \left(\frac{\delta}{70 \binom{K}{m} (K m^2) t^2} \right) \\ &\leq K \left((13m + 1 + (30m^2 - 26m)) \right) \sum_{t=1}^{\infty} \left(\frac{1}{t^2} \right) \sum_{\mathcal{A}} \left(\frac{\delta}{70 \binom{K}{m} (K m^2)} \right) \\ &\leq K \left((13m + 1 + (30m^2 - 26m)) \right) \sum_{\mathcal{A}} \left(\frac{\delta}{70 \binom{K}{m} (K m^2)} \right) \leq \delta. \end{aligned}$$

Now, we show that with probability $1 - \delta$, the best subset can never be eliminated. The best subset A^* gets eliminated if at some time t , for some suboptimal subset A_t , the following condition holds

$$\widehat{\psi}_{A_t} + \alpha_t < \widehat{\psi}_{A^*} - \alpha_t \quad (37)$$

On the event E ,

$$\psi_{A^*} \geq \widehat{\psi}_{A^*} - \alpha_t, \quad \widehat{\psi}_{A_t} + \alpha_t \geq \psi_{A_t} \quad (38)$$

Substituting (38) in (37), we obtain the following: $\psi_{A^*} \geq \widehat{\psi}_{A^*} - \alpha_t \geq \widehat{\psi}_{A_t} + \alpha_t \geq \psi_{A_t}$, and this leads to a contradiction. Hence, w.p. $1 - \delta$, the best subset is never eliminated, and the successive elimination algorithm is $(0, \delta)$ - PAC.

Next, we derive a bound on sample complexity of the successive elimination algorithm.

Notice that, on the event E , from (38), we have $\widehat{\psi}_{A^*} \leq \psi_{A^*} + \alpha_t, \widehat{\psi}_{A_t} \geq \psi_{A_t} - \alpha_t$.

Now, $\widehat{\psi}_{A^*} + \alpha_t \leq \psi_{A^*} + 2\alpha_t \leq \psi_{A_t} - 2\alpha_t \leq \widehat{\psi}_{A_t} - \alpha_t$ which holds if $\psi_{A_t} - \psi_{A^*} \geq 4\alpha_t$ or, equivalently $\Delta(A) \geq 4\alpha_t$.

From (36), we have

$$\alpha_t = \sqrt{a} (c_2 \sqrt{a} + c_1), \quad \text{where } a = \left(\frac{\log \left(\frac{70 \binom{K}{m} K m^2 t^2}{\delta} \right)}{2c_3 t} \right).$$

Solving $\Delta(A) - 4(c_2 a + \sqrt{c_1} \sqrt{a}) \geq 0$, we obtain, as a solution for a , $0 \leq a \leq \frac{-2c_2^2 \sqrt{\frac{c_1(c_1+c_2\Delta(A))}{c_2^4}} + 2c_1 + c_2\Delta(A)}{4c_2^2}$. Therefore,

$a \leq \frac{2c_1+c_2\Delta(A)}{4c_2^2}$. Finally, by solving the equation $\left(\frac{\log \left(\frac{70 \binom{K}{m} K m^2 t^2}{\delta} \right)}{2c_3 t} \right) \leq \frac{2c_1+c_2\Delta(A)}{4c_2^2}$, we obtain

$$t(A) = \mathcal{O} \left(\frac{1}{\Delta(A)} \log \left(\frac{\binom{K}{m} K m^2 \log(\Delta(A)^{-1})}{\delta} \right) \right)$$

where c_4 is a constant. Therefore, w.p. $1 - \delta$, the overall sample complexity is bounded above by

$$\sum_{A \in \mathcal{A}} t(A) = \mathcal{O} \left(\sum_{A \in \mathcal{A}} \frac{1}{\Delta(A)} \log \left(\frac{\binom{K}{m} K m^2 \log(\Delta(A)^{-1})}{\delta} \right) \right).$$

□

D. Lower Bound

Proof. The basis of all the calculations is an established result for the KL-divergence between multivariate Gaussian distributions stated below.

Lemma 3. Let $\mathcal{N}_0, \mathcal{N}_1$ be two k -dimensional normal distribution with zero-mean and covariance matrix A_0, A_1 , respectively,

$$KL(\mathcal{N}_0 || \mathcal{N}_1) = \frac{1}{2} \left[Tr(A_1^{-1} A_0) - k + \ln \left(\frac{\det(A_1)}{\det(A_0)} \right) \right]$$

Using this standard result, we bound KL-divergence between original and transformed problem instances below.

Case $m < j < k$:

When the i th ($i \in \{1, 2\}$) and the m th row of Σ are relabeled, the matrices A_0 and A_1 are $\begin{bmatrix} 1 & \rho^i \\ \rho^i & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & \rho^m \\ \rho^m & 1 \end{bmatrix}$.

Thus,

$$\begin{aligned} KL_{1j}^{1m} &= \frac{1}{2} \left(2 \frac{(1 - \rho^{m+1})}{(1 - \rho^{2m})} - 2 + \ln \left(\frac{1 - \rho^{2m}}{1 - \rho^2} \right) \right) \leq \frac{\rho^2}{2} \left(2 \frac{(\rho^{2m-2} - \rho^{m-1})}{(1 - \rho^{2m})} + \frac{1 - \rho^{2m-2}}{1 - \rho^2} \right) \\ &\leq \frac{\rho^2}{2} \left(\frac{1 - \rho^{2(m-1)}}{1 - \rho^2} \right). \end{aligned}$$

Similarly, $KL_{2j}^{2m} \leq \frac{\rho^4}{2} \left(\frac{1 - \rho^{2(m-2)}}{1 - \rho^4} \right)$, $KL_{mj}^{1m} \leq \frac{\rho^2}{2} \left(\frac{1 - \rho^{m-1}}{1 - \rho^2} \right)$, and $KL_{mj}^{2m} \leq \frac{\rho^4}{2} \left(\frac{1 - \rho^{m-2}}{1 - \rho^4} \right)$. **Case $1 < j < m$:**

$$KL_{1j}^{1m} \leq \frac{\rho^2}{2} \left(\frac{1 - \rho^{2(j-1)}}{1 - \rho^2} \right), KL_{2j}^{2m} \leq \frac{\rho^4}{2} \left(\frac{1 - \rho^{2(j-2)}}{1 - \rho^4} \right), KL_{mj}^{1m} \leq \frac{\rho^2}{2} \left(\frac{1 - \rho^{j-1}}{1 - \rho^2} \right), \text{ and}$$

$$KL_{mj}^{2m} \leq \frac{\rho^4}{2} \left(\frac{1 - \rho^{j-2}}{1 - \rho^4} \right)$$

$KL_{im}^{im} = 0$ for $i \in \{1, 2\}$, $KL_{ij}^{km} = 0$ for $k \in \{1, 2\}$, $i \notin \{1, 2, m\} \cap j \notin \{1, 2, m\}$.

Notice that $\mathbb{E}_{\Sigma}[\tau_{\delta}] = \sum_{(i,j)} \mathbb{E}[N_{ij}(\tau_{\delta})]$. For any $\Sigma' \in \text{Alt}(\Sigma)$, from Lemma 1 and Remark 2 of [14], we have

$$\sum_{(i,j)} \mathbb{E}[N_{ij}(\tau_{\delta})] KL(\Sigma_{X_i X_j} \| \Sigma'_{X_i X_j}) \geq \log \left(\frac{1}{2.4\delta} \right).$$

Consider the following optimization problem, with $\alpha_{ij} \triangleq \mathbb{E}[N_{ij}(\tau_{\delta})]$:

$$\min_{\{\alpha_{ij}\}} \sum_{(i,j)} \alpha_{ij} \quad \text{subject to} \quad \mathbb{E}[\tau_{\delta}] \sum_{(i,j)} \frac{\alpha_{ij}}{\mathbb{E}[\tau_{\delta}]} KL(\Sigma_{X_i X_j} \| \Sigma'_{X_i X_j}) \geq \log \left(\frac{1}{2.4\delta} \right), \forall \Sigma' \in \text{Alt}(\Sigma).$$

Letting $w_{ij} \triangleq \frac{\alpha_{ij}}{\mathbb{E}[\tau_{\delta}]}$, the problem defined above is equivalent to the following:

$$\min_{\{w_{ij}\}} \mathbb{E}[\tau_{\delta}] \quad \text{subject to} \quad \mathbb{E}[\tau_{\delta}] \min_{\Sigma' \in \text{Alt}(\Sigma)} \sum_{(i,j)} w_{ij} KL(\Sigma_{X_i X_j} \| \Sigma'_{X_i X_j}) \geq \log \left[\frac{1}{2.4\delta} \right].$$

Hence,

$$\min_{\{w_{ij}\}} \mathbb{E}[\tau_{\delta}] \geq \frac{\log(\frac{1}{2.4\delta})}{\max_{w \in \Delta_{\binom{K}{2}}} \min_{\Sigma' \in \text{Alt}(\Sigma)} \sum_{(i,j)} w_{ij} KL(\Sigma_{X_i X_j} \| \Sigma'_{X_i X_j})}.$$

Next, we derive an upper bound on the max-min in the denominator above.

Let $f(w, \Sigma') \triangleq \sum_{(i,j)} w_{ij} KL(\Sigma_{X_i X_j} \| \Sigma'_{X_i X_j})$. Then, we have

$$\begin{aligned} f(w, \Sigma^{13}) &= w_{12} KL_{12}^{13} + w_{14} KL_{14}^{13} + \dots + w_{1K} KL_{1K}^{13} + w_{23} KL_{23}^{13} + w_{34} KL_{34}^{13} + \dots + w_{3K} KL_{3K}^{13} \\ &\leq \frac{\rho^2}{2(1 - \rho^2)} (w_{12}(1 - \rho^2) + (w_{14} + \dots + w_{1K})(1 - \rho^4) + w_{23}(1 - \rho) + (w_{34} + \dots + w_{3K})(1 - \rho^2)), \\ &\vdots \\ f(w, \Sigma^{1m}) &= \sum_{\substack{j=2 \\ j \neq m}}^K (w_{1j} KL_{1j}^{1m} + w_{mj} KL_{mj}^{1m}) \leq \frac{\rho^2}{2(1 - \rho^2)} \left(\sum_{2 \leq j < m} w_{1j} (1 - \rho^{2(j-1)}) + \right. \\ &\quad \left. \sum_{m < j \leq K} w_{1j} (1 - \rho^{2(m-1)}) + \sum_{2 \leq j < m} w_{mj} (1 - \rho^{j-1}) + \sum_{m < j \leq K} w_{mj} (1 - \rho^{m-1}) \right). \end{aligned}$$

Along similar lines,

$$\begin{aligned} f(w, \Sigma^{23}) &\leq \frac{\rho^4}{2(1 - \rho^4)} ((1 - \rho^2)(w_{24} + \dots + w_{2K}) + (1 - \rho)(w_{34} + \dots + w_{3K})), \\ &\vdots \\ f(w, \Sigma^{2m}) &= \sum_{\substack{j=3 \\ j \neq m}}^K (w_{2j} KL_{2j}^{2m} + w_{mj} KL_{mj}^{2m}) \leq \frac{\rho^4}{2(1 - \rho^4)} \left(\sum_{3 \leq j < m} w_{2j} (1 - \rho^{2(j-2)}) + \right. \\ &\quad \left. \sum_{m < j \leq K} w_{2j} (1 - \rho^{2(m-2)}) + \sum_{3 \leq j < m} w_{mj} (1 - \rho^{j-2}) + \sum_{m < j \leq K} w_{mj} (1 - \rho^{m-2}) \right). \end{aligned}$$

Notice that $\min_w f(w, \Sigma^{1m}) \geq \max_w f(w, \Sigma^{2m})$. This inequality holds because $\rho^2(1 - \rho^{2(m-2)}) \geq (1 + \rho^2)(1 - \rho)$ for $m \geq 3$, and $\rho \in [-1, 1]$.

Now,

$$\begin{aligned}
& \max_w \min (f(w, \Sigma^{23}), f(w, \Sigma^{24}), \dots, f(w, \Sigma^{2K})) \\
&= \max_w \min_{\substack{\alpha_1, \alpha_2, \dots, \alpha_K \geq 0 \\ \alpha_1 + \alpha_2 + \dots + \alpha_K = 1}} \{ \alpha_1 f(w, \Sigma^{23}) + \alpha_2 f(w, \Sigma^{24}) + \dots + \alpha_K f(w, \Sigma^{2K}) \} \\
&\leq \min_{\substack{\alpha_1, \alpha_2, \dots, \alpha_K \geq 0 \\ \alpha_1 + \alpha_2 + \dots + \alpha_K = 1}} \max_w \{ \alpha_1 f(w, \Sigma^{23}) + \alpha_2 f(w, \Sigma^{24}) + \dots + \alpha_K f(w, \Sigma^{2K}) \} \\
&\leq \max_w f(w, \Sigma^{23}) \text{ (choosing } \alpha_1 = 1, \alpha_i = 0, i = 2, \dots, K) \\
&= \max_w \left\{ \frac{\rho^4}{2(1-\rho^4)} ((1-\rho^2)(w_{24} + \dots + w_{2K}) + (1-\rho)(w_{34} + \dots + w_{3K})) \right\} = \frac{\rho^4}{2(1+\rho^2)},
\end{aligned}$$

where the final inequality holds for any $\rho \in [0, 1]$, with the following optimal weights: $\sum_{j=4}^K w_{2j} = 1$, and $\sum_{j=4}^K w_{3j} = 0$.

Notice that the smallest gap Δ for the bandit instance governed by Σ is given by

$$\Delta = \psi(\{2, 3\}) - \psi(\{1, 2\}) = \frac{1}{(1-\rho^4)} [(K-3)(\rho^2 + 3\rho^4 + 2\rho^6 - 2\rho^7) + (2\rho^4 + 3\rho^6 - \rho^2)].$$

A simple calculation yields $\frac{\rho^4}{4(1+\rho^2)} \leq \Delta$, which implies $\mathbb{E}[\tau_\delta] \geq \frac{\log(\frac{1}{2.4\delta})}{\Delta}$. Hence proved. \square