Harnessing Structure in Discrete and Non-Convex Optimization

with Applications in Online Learning, Multi-Agent Systems, and Phase Retrieval

by

Parth Kashyap Thaker

A Dissertation Presented in Partial Fulfillment
of the Requirement for the Degree
Doctor of Philosophy

Approved December 2023 by the
Graduate Supervisory Committee:

Gautam Dasarathy, Chair
Lalitha Sankar
Angelia Nedić
Nicolò Michelusi

ARIZONA STATE UNIVERSITY

May 2024

ABSTRACT

This thesis examines the critical relationship between data, complex models, and other methods to measure and analyze them. As models grow larger and more intricate, they require more data, making it vital to use that data effectively. The document starts with a deep dive into nonconvex functions, a fundamental element of modern complex systems, identifying key conditions that ensure these systems can be analyzed efficiently—a crucial consideration in an era of vast amounts of variables.

Loss functions, traditionally seen as mere optimization tools, are analyzed and recast as measures of how accurately a model reflects reality. This redefined perspective permits the refinement of data-sourcing strategies for a better data economy. The aim of the investigation is the model itself, which is used to understand and harness the underlying patterns of complex systems. By incorporating structure both implicitly (through periodic patterns) and explicitly (using graphs), the model's ability to make sense of the data is enhanced.

Moreover, online learning principles are applied to a crucial practical scenario: robotic resource monitoring. The results established in this thesis, backed by simulations and theoretical proofs, highlight the advantages of online learning methods over traditional ones commonly used in robotics.

In sum, this thesis presents an integrated approach to measuring complex systems, providing new insights and methods that push forward the capabilities of machine learning.

To my Father,

May you find eternal peace.

ACKNOWLEDGEMENT

First and foremost, I would like to thank my advisor Prof. Gautam Dasarathy for supporting me through thick and thin. Graduating with a Ph.D. wouldn't have been possible without him. I would like to also thank Prof. Angelia Nedić for giving me a chance to pursue this ardent dream of mine. I also want to thank Prof. Lalitha Sankar and Prof. Nicolò Michelusi for their valuable input.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

This thesis examines the critical relationship between data, complex models, and the methods we use to measure and analyze them. As models grow larger and more intricate, they require more data, making it vital to use that data effectively. We begin with a deep dive into nonconvex functions, a fundamental element of modern complex systems, identifying key conditions that ensure these systems can be analyzed efficiently—a crucial consideration in an era of vast amounts of variables. Nonconvex optimization problems are notoriously difficult to solve due to the presence of multiple local minima, saddle points, and maxima, which complicate the search for global solutions. This challenge is directly addressed in Chapter 3. This research contributes significantly to the understanding of nonconvex optimization, paving the way for the development of more effective strategies for dealing with such problems in various applications.

We then take a closer look at loss functions, traditionally seen as mere optimization tools, and recast them as measures of how accurately a model reflects reality. This redefined perspective allows us to refine data-sourcing strategies for a better data efficiency. The crux of our investigation is the model itself, which we use to understand and harness the underlying patterns of complex systems. By incorporating structure both implicitly (through periodic patterns) and explicitly (using graphs), we enhance the model's ability to make sense of the data. This aspect is tackled through the implementation of bandit algorithms and the use of graph theory as a regularization tool (Chapter 2) and periodic functions as implicit regularizers (Chapter 5). Through these works, we demonstrate how bandit algorithms can lead to more efficient data

generation. Furthermore, the incorporation of graph structures and the exploitation of periodic functions as implicit structures contribute to making data generation more efficient, reducing the overall data requirements for complex models.

The third challenge this thesis addresses is the need for provable and efficient implementation methods for these increasingly complex models. This is a critical area, as the practical applicability of theoretical models depends heavily on their implementation efficiency and reliability. In tackling this, Chapter 3 provides a foundation by proving the efficacy of first-order methods in nonconvex settings. Moreover, Chapter 4 extends this theme by developing algorithms that enable drones to prioritize data generation for important phenomena, showcasing a practical application where efficient data collection is paramount. This approach not only enhances the efficiency of data collection but also ensures the reliability and applicability of the models in real-world scenarios.

In sum, this work presents an integrated approach to measuring complex systems, providing new insights and methods that push forward the capabilities of machine learning.

Chapter 2

MAXIMIZING AND SATISFICING IN MULTI-ARMED BANDITS WITH GRAPH INFORMATION

2.1 Introduction

The multi-armed bandit has emerged as an important paradigm for modeling sequential decision making and learning under uncertainty. Practical applications include design policies for sequential experiments Robbins (1952), combinatorial online leaning tasks Chen *et al.* (2014), collaborative learning on social media networks Kolla *et al.* (2018); Audibert *et al.* (2010), latency reduction in cloud systems Joshi *et al.* (2017) and many others Cao *et al.* (2015); Zhou *et al.* (2014); Tekin and Turğay (2018); Kandasamy *et al.* (2016). In the traditional multi-armed bandit problem, the goal of the agent is to sequentially choose among a set of actions or arms to maximize a desired performance criterion or reward. This objective demands a delicate tradeoff between exploration (of new arms) and exploitation (of promising arms). An important variant of the reward maximization problem is the identification of arms with the highest (or near-highest) expected reward. This *best arm identification* Mannor and Tsitsiklis (2004); Even-Dar *et al.* (2006) problem, which is one of pure exploration, has a wide range of important applications like identifying and testing drugs to treat infectious diseases like COVID-19, finding relevant users to run targeted ad campaigns, hyperparameter optimization in neural networks and recommendation systems. The broad range of applications of this paradigm is unsurprising given its ability to essentially model any optimization problem of black-box functions on discrete (or discretizable) domains with noisy observations.

While pure exploration problems in bandits show considerable promise, there are significant hurdles to their practical usage. In modern applications, one is often faced with a tremendously large number of options (sometimes in the order of hundreds of millions) that need to be considered for decision making. In such cases, playing (i.e., obtaining a random sample from) each bandit arm even once could be intractable. This renders traditional approaches to pure exploration ineffective. Fortunately, in several applications, the arms and their rewards are related to each other and information about the reward of one arm may be deduced from plays of similar arms. In this paper, we consider the pure exploration problem in stochastic multi-armed bandits where the similarities between arms are captured by a graph and the rewards may be represented as a smooth signal on this graph. Such graph side information is available in a wide range of applications: search and recommendation systems have graphs that capture similarities between items (Guo *et al.*, 2010; Rao *et al.*, 2015; Wu *et al.*, 2020; Dasarathy *et al.*, 2017); drugs, molecules and their interactions can be represented on a graph (Ioannidis *et al.*, 2020); targeted advertising considers users connected to each other in a social network (Jamali and Ester, 2009), and hyperparameters for training neural network are often inter-related (Young *et al.*, 2018). It is worth noting that such graphs are sometimes intrinsic to the problem (e.g., spatial coordinates or social/computer networks), or may be inferred based on similarity metrics defined on arm features; a recent line of work considers constructing such graphs to enable more effective learning (see e.g., Zhang *et al.*, 2022; Kushnir and Venturi, 2020).

2.1.1 Related Work

The textbook Lattimore and Szepesvári (2020) is an excellent resource for the general problem of multi-armed bandits. The pure exploration variant of the bandit problem is more recent and has also received considerable attention in the litera-

ture Bubeck *et al.* (2009, 2011); Garivier and Kaufmann (2019); Gabillon *et al.* (2012); Audibert *et al.* (2010); Jamieson and Nowak (2014). These lines of work treat the bandit arms or actions as independent entities and playing a particular arm yields no information about any other arm. This leads to great difficulty in scaling such methods, since in the problem setups with a large number of arms, attempting to play *all* arms is not practical. We resolve this precise roadblock by introducing a convenient way of appending graph side information into the mix which provably accelerates the process of sub-optimal arm elimination (potentially without playing it even once!)

A recent line of work Li *et al.* (2016); Lattimore and Szepesvári (2020); Gupta *et al.* (2020); Yang *et al.* (2020); Gentile *et al.* (2014); Ma *et al.* (2015a) has proposed the leveraging of structural side-information for the multi-armed bandit problem for regret minimization. Such topology-based bandit methods work under the assumption that pulling an arm reveals information about other, correlated arms Gupta *et al.* (2020); Shamir (2011), which helps in developing better regret methods. Similarly, spectral bandits Kocák and Garivier (2020); Yang *et al.* (2020); Valko *et al.* (2014) assume user features are modeled as signals defined on an underlying graph, and use this to assist in learning. The works Atsidakou *et al.* (2022) and Wu *et al.* (2015) consider similar graph information models, albeit at a degraded level. The authors in LeJeune *et al.* (2020) use the graphs to improve the regret bounds in a thresholding bandit setting. Work revolving around spectral bandits utilizes the *spectrum* of the graph laplacian. In contrast, we focus on the *combinatorial properties* of the graphs to devise algorithms and analyze them. Another line of work Dasarathy *et al.* (2015); Wang *et al.* (2019); Lipor and Dasarathy (2018); Ma *et al.* (2013a) considers search problems on graphs under a different model and there is an opportunity for future work to combine these techniques.

Most of the aforementioned works focus on regret minimization in the presence

of graph information. The problem of pure exploration with similarity graphs has received far less attention. The authors in Kocák and Garivier (2020) were the first to attempt to fill this gap for the spectral bandit setting. They provide an information-theoretic lower bound and a gradient-based algorithm to estimate this lower bound to sample the arms. The authors provide performance guarantees for the algorithm, but these results only indirectly capture the benefit brought by the graph; our results on the other hand are based on a novel complexity measure that explicitly elicits the benefit of having the graph side information.

Note that, similarity graph information considered in this work is fundamentally different from linear rewards assumption in contextual/linear bandits. In the linear bandits problem, the reward behavior is assumed to be low dimensional and this is crucial for the improved regret bounds and sample complexity guarantees Lattimore and Szepesvári (2020); Soare *et al.* (2014). In the current work, we do not make any assumptions on the low dimensionality of the rewards but still show improvements in sample complexity provided a good arm-similarity graph is available. We show a toy example in Appendix 2.8 where a low dimensional linear bandit cannot be competitive with the corresponding graph-bandit setting.

2.1.2 Overview of Contributions

We consider the pure exploration of multi-arm bandits problem when a graph that captures similarities between the arms is available. In particular, we consider the problem of finding the arm with the maximum reward (i.e., the maximizing problem) or one that has sufficiently high reward (i.e., the satisficing problem [1]) under the assumption that arm rewards are smooth with respect to a known graph. Our main contributions may be summarized as follows:

[1] named after Herbert Simon's celebrated alternative model of decision making Simon (1955)

(a) We devise a novel algorithm GRUB for the best arm identification problem (i.e., the maximizing problem) that specifically exploits the *homophily* (strong connections imply similar average rewards) on the graph (Section 2.3).

(b) We provide a theoretical characterization of the performance of GRUB. To this end, we define a novel measure $\mathfrak{I}$ that we dub the "*influence factor*" which depends on the resistance distance of the underlying graph. This measure captures the benefit of the graph side information and plays a central role in the analysis of GRUB. In the traditional (graph-free) best arm identification problem, the sample complexity is known to scale as $\sum_{i=1}^{n} \frac{1}{\Delta_i^2}$, where Δ_i is the gap between the expected rewards of the best arm and arm i. On the other hand, we show that GRUB roughly has a complexity that scales like $\sum_{i\in\mathcal{H}} \frac{1}{\Delta_i^2}$ samples where the set $\mathcal{H}$ is a set dependent on the influence factor, which contains arms which are hard to distinguish from the optimal arm. For a broad range of problems $|\mathcal{H}| \ll n$, yielding significant improvement over traditional best arm identification algorithms (Section 2.4).

(c) In Section 2.5, we provide lower bounds on the minimum number of samples required for the identification of the optimal arm when a graph encoding arm similarities is available. This shows the near-optimality of GRUB for an important class of representative problems.

(d) In many real-world scenarios, the aim of finding the absolute best arm can often be too costly or even intractable. In these situations, it may be more appropriate to solve the *satisficing* problem, where the algorithm returns an arm that is good enough. We propose a variant of GRUB, dubbed ζ-GRUB for this important setting in Section 2.6

(e) Finally, in Section 5.6, we complement our theoretical results with an empirical evaluation of our algorithms. We further provide algorithmic improvements to GRUB and discuss novel sampling policies for best arm identification in the presence of graph

information.

2.2 Problem Setup and Notation

We consider an n-armed bandit problem with the set of arms given by $[n] \triangleq \{1, 2, 3, \ldots, n\}$. Each arm $i \in [n]$ is associated with a σ-sub-Gaussian distribution ν_i. That is, $\mathbb{E}_{X \sim \nu_i}\left[\exp\left(s(X - \mu_i)\right)\right] \leq \exp\left(\frac{\sigma^2 s^2}{2}\right)\ \forall s \in \mathbb{R}$, where $\mu_i = \mathbb{E}_{\nu_i}[X]$ is said to be the (expected or mean) reward associated to arm i. We will let $\boldsymbol{\mu} \in \mathbb{R}^n$ denote the vector of all the arm rewards. A "play" of an arm i is simply an observation of an independent sample from ν_i; this can be thought of as a noisy observation of the corresponding mean μ_i. The goal of the best-arm identification problem is to identify, from such noisy samples, the arm $a^* \triangleq \arg\max_{i \in [n]} \mu_i$ that has the maximum expected reward, denoted by μ^*. For each arm $i \in [n]$, we will let $\Delta_i \triangleq \mu^* - \mu_i$ denote the sub-optimality of the arm.

As discussed in Section 5.2, our goal is to consider the best-arm identification where one has additional access to information about the similarity of the arms under consideration. In particular, we model this side information as a weighted undirected graph $G = (V_G, E_G, A_G)$ where the vertex set, $V_G = [n]$, is identified with the set of arms, the edge set $E_G \subseteq \binom{[n]}{2}$, and adjacency matrix $A_G \in \mathbb{R}^{n \times n}$ describes the weights of the edges E between the arms which capture the similarity in means of connected arms; the higher the weight, the more similar the rewards from the corresponding arms. We will let $L_G = D_G - A_G$ denote the combinatorial Laplacian [2] of the graph Chung and Graham (1997), where $D_G = \mathrm{diag}(A_G \times \mathbb{1}_n)$ is a diagonal matrix containing the weighted degrees of the vertices. We will suppress the dependence on G when the context is clear. Subsequently, we show that if one has access to this graph and

[2] All our results continue to hold if this is replaced with the normalized, random walk, or generalized Laplacian.

the vector of rewards $\boldsymbol{\mu}$ is *smooth* with respect to the graph (that is, highly similar arms have highly similar rewards), then one can solve the pure exploration problem extremely efficiently. We will capture the degree of smoothness of $\boldsymbol{\mu}$ with respect to the graph using the following seminorm [3] :

$$\|\boldsymbol{\mu}\|_G^2 \triangleq \langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle = \sum_{\{i,j\} \in E_G} A_{ij}(\mu_i - \mu_j)^2. \tag{2.1}$$

The second equality above can be verified by a straightforward calculation. Also, notice that $\|\boldsymbol{\mu}\|_G$ being small implies $\mu_i \approx \mu_j$ for $(i,j) \in E$. In such scenario, we say that the mean vector $\boldsymbol{\mu}$ is smooth over graph G. This observation has inspired the use of the Laplacian in several lines of work to enforce smoothness on the vertex-valued functions Ando and Zhang (2007); Valko *et al.* (2014); Zhu (2005); LeJeune *et al.* (2020). For $\epsilon > 0$, we say that arms (rewards) are ϵ-smooth with respect to a graph G if $\|\boldsymbol{\mu}\|_G \leq \epsilon$.

Let $\mathcal{C}(G) \subset 2^{[n]}$ denote the set of all connected components and let $k(G) \triangleq |\mathcal{C}(G)|$ denote the number of connected components of the graph G. For a vertex $i \in [n]$, we will let $C_i(G) \in \mathcal{C}(G)$ denote the connected component that contains i. When the context is clear we sometimes let $C_i(G)$ also refer all the nodes in the connected component. We say a graph $G = ([n], E)$ has *k-isolated cliques* if it can be divided into fully connected sub-graphs $G_i = (V_i, E_i)$ such that $V_i \subseteq [n], E_i = \binom{V_i}{2}$ for all $i \in [k], V_i \cap V_j = \emptyset, E_i \cap E_j = \emptyset$ for all $i, j \in [k]$, and $\bigcup_{i=1}^k V_i = [n], \bigcup_{i=1}^k E_i = E$. Notice that we only have one clique if G is fully connected.

To solve the best-arm identification problem, we need a sampling policy to sequentially and interactively select the next arm to play, and a stopping criterion. For any time $t \in \mathbb{N}$, the sampling policy $\boldsymbol{\pi}_t = \{\pi_s\}_{s \leq t}$ is a function that maps t to an arm in $[n]$ given the history of observations up to time $t-1$. With slight abuse of notation, we

[3] L_G is not positive definite and can be verified to have as many zero eigenvalues as the number of connected components in G

will let π_t denote the arm chosen by an agent at time t. Let r_{t,π_t} denote the random reward observed at time t from arm π_t. We use $t_i(\boldsymbol{\pi}_t)$ (referred as t_i for simplicity) to denote the number of times arm i is played under the sampling policy $\boldsymbol{\pi}_t$. In this paper, we tackle the following problems:

P1 (Best arm identification): *Given n arms and an arbitrary graph G capturing similarity between the arms, can we design a policy $\boldsymbol{\pi}_T$ that exploits the similarity to find the best arm efficiently?*

P2 (ζ-best arm identification): *Under the setting in **P1**, can we design a similarity exploiting policy $\boldsymbol{\pi}_T$ so as to find an arm belonging to the set $B(\zeta) \triangleq \{i \in [n] : |\mu_i - \mu_{a^*}| \leq \zeta\}$ efficiently?*

2.3 The GRUBAlgorithm

We now introduce GRUB (GRaph based Upper Confidence Bound), a novel but natural algorithm for best arm identification in the presence of graph side information. We begin with an intuitive description of how GRUB incorporates the graph side information into an *upper confidence bound* (UCB) strategy. Most UCB algorithms Lattimore and Szepesvári (2020); Valko *et al.* (2014) compute the estimates of mean and variance, and use these to eliminate arms that have been deduced to be sub-optima. The key idea behind GRUB is that the arm similarity information allows us to create high-quality estimates of mean rewards and confidence intervals for arms that have not been (sufficiently) sampled yet. In what follows, we describe the building blocks of GRUB.

2.3.1 Leveraging Graph Side Information

We introduce two key ideas that lie at the heart of the GRUB algorithm. First, at each step, GRUB computes a regularized estimate of the means of *all the arms*; the regularization based on the graph Laplacian essentially promotes the smoothness of the mean vector on the given graph. This allows the algorithm to estimate the means of arms it has *never sampled*. To do this, at any given time step T, the algorithm solves the following Laplacian-regularized least-squares optimization program:

$$\hat{\boldsymbol{\mu}}_T = \underset{\boldsymbol{\mu}\in\mathbb{R}^n}{\arg\min} \left\{ \left[\sum_{t=1}^{T}(r_{t,\pi_t} - \mu_{\pi_t})^2\right] + \rho\langle\boldsymbol{\mu}, L_G\boldsymbol{\mu}\rangle \right\}, \tag{2.2}$$

where $\rho > 0$ is a tunable parameter. Equation (2.2) admits a closed form solution of the form

$$\hat{\boldsymbol{\mu}}_T = \left(\sum_{t=1}^{T}\mathbf{e}_{\pi_t}\mathbf{e}_{\pi_t}^\top + \rho L_G\right)^{-1}\left(\sum_{t=1}^{T}\mathbf{e}_{\pi_t}r_{t,\pi_t}\right),$$

provided the matrix $V_T \triangleq \sum_{t=1}^{T}\mathbf{e}_{\pi_t}\mathbf{e}_{\pi_t}^\top + \rho L_G$ is invertible; $\mathbf{e}_i$ denotes the i-th standard basis vector for the Euclidean space $\mathbb{R}^n$. In Appendix A.1 we show that invertibility holds if and only if the sampling policy yields at least one sample per connected component of G. This is a rather mild condition that we arrange explicitly in our algorithm, given that we know the graph G. In what follows we assume that every connected component of graph G is sampled at least once. This regularized mean estimation procedure yields an estimate of the mean that is both in agreement with observations and smooth on the graph – thereby allowing information sharing among similar arms.

The second key idea of our algorithm is the utilization of the graph G in tracking the confidence bounds of *all the arms simultaneously*. Intuitively, for identifying the best arm, we must be reasonably certain about the sub-optimality of the other arms. This in turn would require the algorithm to track a high-probability confidence bound

on the means of all the arms. In the traditional (graph-free) best arm identification problem, the confidence interval of an arm's mean estimate depends on the number of times the arm has been played. Requiring multiple plays of all suboptimal arms for obtaining high confidence bounds is potentially disastrous when the number of arms is very large. In our setup, we show that the knowledge of the similarity graph greatly improves this situation. In particular, we show that a play of any arm not only tightens its own confidence interval but also has an impact on the confidence intervals of *all connected arms*. To quantify the benefit of graph information for the confidence bounds, we will define a novel quantity for each arm – the effective number of plays.

Definition 2.3.1 (Effective Number of Plays)**.** *Let $\rho > 0$ and $\{t_i\}_{i=1}^n$ denote the number of plays of each of the n arms when a sampling policy $\boldsymbol{\pi}_T$ is employed for T time steps. Suppose that for each connected component $C \in \mathcal{C}(G)$, there is at least one arm $i_C \in C$ such that $t_{i_C} > 0$. Then the effective number of plays for each arm $i \in [n]$ is defined as $t_{\mathit{eff},i} \triangleq \left[(N_T + \rho L_G)^{-1}\right]_{ii}^{-1}$, where N_T is a diagonal matrix of $\{t_i\}_{i=1}^n$, and L_G denotes the Laplacian of the given graph G.*

The effective number of plays $t_{\mathrm{eff},i}$ for any arm i is influenced by two factors: (a) the number of samples of arm i itself, and (b) the number of samples of any arm in the connected component $j \in C(i), j \neq i$. It can be shown that for any arm i, $t_{\mathrm{eff},i}$ depends on the number of connections of node i in graph G, and its value increases as the connectivity of the node increases. The choice of the terminology for this quantity is justified by the following lemma, which provides a high confidence bound for the mean estimate of each arm.

Lemma 2.3.2 (Concentration inequality)**.** *For any $T > k(G)$, the following holds*

with probability at least $1-\delta$*:*

$$|\hat{\mu}_T^i - \mu_i| \leq \sqrt{\frac{1}{t_{eff,i}}}\left(2\sigma\sqrt{14\log\left(\frac{2w_i(\boldsymbol{\pi}_T)}{\delta}\right)} + \rho\|\boldsymbol{\mu}\|_G\right), \quad \forall i \in [n] \tag{2.3}$$

where $w_i(\boldsymbol{\pi}_T) = a_0 n t_{eff,i}^2$ *for any constant* $a_0 > 0$*,* $\hat{\mu}_T^i$ *is the* i*-th coordinate of the estimate from* (2.2)

Notice that the *effective number of plays* has a similar role as the number of plays in traditional pure exploration algorithms Even-Dar *et al.* (2006). Indeed, in the absence of graph information, $t_{\text{eff},i}$ reduces to t_i, the total number of plays of individual arms. Lemma 2.3.2 recovers high confidence bounds for standard best-arm identification problem Even-Dar *et al.* (2006). It should be noted that while our work is the first to identify this interpretable quantity explicitly, the result of Lemma 2.3.2 in other forms has appeared before in the literature Abbasi-yadkori *et al.* (2011); Valko *et al.* (2014); Yang *et al.* (2020).

We introduce our algorithm GRUB for best arm identification when the arms can be approximately cast as nodes on a graph. GRUB uses insights from graph-based mean estimation (2.2) and upper confidence bound estimation (2.3) for its elimination policies to search for the optimal arm.

GRUB accepts as input a graph G on n arms (and its Laplacian L_G), a regularization parameter $\rho > 0$, a smoothness parameter $\epsilon > 0$, and an error tolerance parameter $\delta \in (0, 1)$. It is composed of the following major blocks.

Initialization: First, GRUB identifies the clusters in the G using a `Cluster-Identification` routine. Any algorithm that can efficiently partition a graph can be used here, e.g. METIS Karypis and Kumar (1998). GRUB then samples one arm from each cluster. This ensures $V_T \succ 0$, which enables GRUB to estimate $\hat{\boldsymbol{\mu}}_T$ using the closed form solution of eq. (2.2). A great advantage of GRUB is that the initialization phase only requires steps equal to the number of disconnected components in the graph. This is

in direct contrast with traditional best arm identification algorithms, which require atleast one sample from every arm initially.

Sampling policy: At each round, GRUB obtains a sample from the arm returned by the routine `Sampling-Policy`, which cyclically samples arms from different clusters while ensuring that no arm is resampled before all arms in consideration have the same number of samples. This is distinct from standard cyclic sampling policies that is traditionally used for best arm identification Even-Dar *et al.* (2006), but any of them may be modified readily to provide a cluster-aware sampling policy for GRUB. In our experiments, we show that replacing cyclic sampling with more statistics- and structure-aware sampling greatly improves performance; a theoretical analysis of these is a promising avenue for future work. One of the major advantages of GRUB is the lite nature of the computation. Every loop just requires a rank-1 inverse update which can be performed very efficiently and it does not need any subroutines, unlike (Kocák and Garivier, 2020)

Bad arm elimination : At any time t, let A be the set of all arms in consideration for being optimal. Using the uncertainty bound from (2.3), GRUB uses the following criteria for sub-optimal arm elimination. At each iteration, GRUB identifies an arm $a_{\max} \in A$, $a_{\max} = \arg\max_{i \in A} \left[\hat{\mu}_t^i - \beta_i(t)\sqrt{t_{\text{eff},i}^{-1}}\right]$, where $\beta_i(t) = \left(2\sigma\sqrt{14\log\left(\frac{2na_0t_{\text{eff},i}^2}{\delta}\right)} + \rho\epsilon\right)$, with the *highest lower bound* on its mean estimate.Following this, GRUB removes arms from the set A according to the following elimination policy,

$$A \leftarrow \left\{\mathbf{a} \in A \mid \hat{\mu}_t^{a_{\max}} - \hat{\mu}_t^a \leq \beta_a(t)\sqrt{t_{\text{eff},a}^{-1}} + \beta_{a_{\max}}(t)\sqrt{t_{\text{eff},a_{\max}}^{-1}}\right\}. \tag{2.4}$$

Note that GRUB does not require any optimization innerloop as in Kocák and Garivier (2020). This potentially provides GRUB with a significant computation advantage, especially when the dimensionality of the problem is very large. The pseudocode for GRUB can be found in Appendix A.4.

Algorithm 1 GRUB

Input: Regularization parameter ρ, Smoothness parameter ϵ, Error bound δ, Total arms n, Laplacian L_G, Sub-gaussianity parameter σ

$t \leftarrow 0$

$A = \{1, 2, \ldots, n\}$

$t = 0$

$V_0 \leftarrow \rho L_G$

$\mathcal{C}(G) \leftarrow$ `Cluster-Identification`(L_G)

for $C \in \mathcal{C}(G)$ **do**

  $t \leftarrow t + 1$

  Pick random arm $k \in C$ to observe reward $r_{t,k}$

  $V_t \leftarrow V_{t-1} + \mathbf{e}_k \mathbf{e}_k^T$, and $\mathbf{x}_t \leftarrow \mathbf{x}_{t-1} + r_{t,k}\mathbf{e}_k$

end for

while $|A| > 1$ **do**

  $t \leftarrow t + 1$

  for $i \in A$ **do**

    $t_{\text{eff},i} \leftarrow ([V_t^{-1}]_{ii})^{-1}$

    $\beta_i(t) \leftarrow 2\sigma\sqrt{14 \log\left(\frac{2nt_{\text{eff},i}^2}{\delta}\right)} + \rho\epsilon$

  end for

  $k \leftarrow$ `Sampling-Policy`$(t, V_t, A, \mathcal{C}(G))$

  Sample arm k to observe reward $r_{t,k}$

  $V_t \leftarrow V_{t-1} + \mathbf{e}_k \mathbf{e}_k^T$

  $\mathbf{x}_t \leftarrow \mathbf{x}_{t-1} + r_{t,k}\mathbf{e}_k$

  $\hat{\boldsymbol{\mu}}_t \leftarrow V_t^{-1}\mathbf{x}_t$

  $a_{\max} \leftarrow \arg\max_{i \in A} \left[\hat{\mu}_t^i - \beta(t)\sqrt{t_{\text{eff},i}^{-1}}\right]$

  $A \leftarrow \Big\{\mathbf{a} \in A \mid \hat{\mu}_t^{a_{\max}} - \hat{\mu}_t^a \leq \beta_a(t)\sqrt{t_{\text{eff},a}^{-1}}$

    $+\beta_{a_{\max}}(t)\sqrt{t_{\text{eff},a_{\max}}^{-1}}\Big\}$

end while

return A

Next, we derive performance guarantees on the sample complexity for GRUB to return the best arm with high probability.

2.4 Theoretical Analysis of GRUB

In this section, we provide a formal statement of the sample complexity of GRUB. To do this, we first introduce a novel quantity we call *influence factor*. The influence factor of an arm is derived from resistance distance, a classical graph theoretic concept. This adds to the interpretability and understanding of the instances where using graph side information might be of tremendous use to the application. The usage of graphs through the influence factor allows us to identify arms that can be eliminated quickly from consideration.

2.4.1 Resistance Distance and Influence Factor

We first recall the definition of resistance distance in a graph.

Definition 2.4.1 (Resistance Distance)**.** *Bapat and Gupta (2010) For any graph G with n nodes, given a constant $\delta > 0$, the* ***resistance distance*** *$r_{\delta,G}(i,j)$ between two nodes i, j is defined as,*

$$r_{\delta,G}(i,j) = R_{ii} + R_{jj} - R_{ij} - R_{ji}, \tag{2.5}$$

where $R \triangleq \left(L_G + \delta \mathbb{1}\mathbb{1}^T\right)^\dagger$; $\dagger$ denotes the Moore-Penrose inverse, L_G is the Laplacian of graph G, and $\mathbb{1} \in \mathbb{R}^n$ is the vector of all 1's.

When the context is clear we denote the resistance distance simply as $r_G(\cdot,\cdot)$. The terminology comes from circuit theory: Suppose that a graph $G = ([n], E)$ is thought of as a resistor network on the nodes $[n]$ where each edge $\{i, j\}$ has a unit resistance. Then, the effective resistance between two nodes i and j is precisely the resistance distance $r(i, j)$. It can be shown in general that nodes that are close by or connected

by several paths have a small resistance distance. Given its ability to capture the closeness of nodes in the graph, the resistance distance has found a broad range of applications and has been the subject of many studies; see e.g., (Klein and Randić, 1993; Bapat and Gupta, 2010; Xiao and Gutman, 2003).

Using the notion of resistance distance, we define the influence factor $\mathfrak{I}(\cdot, G)$ of a vertex below. This novel measure quantifies the impact of the graph on the parameter estimation of arm j, and in particular, allows us to use the combinatorial properties of the graph and the arm means to classify arms into two sets: competitive and non-competitive; the definition of these sets follows right after. As our theory will show, the competitive arms are sampled as though we were in the traditional graph-free setting; on the other hand, non-competitive arms are eliminated rapidly, often with zero plays! Indeed, the smoother the reward vector is with respect to the graph, the fewer competitive arms there are – it is this phenomenon that is captured using the influence factor.

Definition 2.4.2 (Influence Factor)**.** *Let G be a graph on the vertex set $[n]$. For each $j \in [n]$, define **influence factor** $\mathfrak{I}(j, G)$ as:*

$$\mathfrak{I}(j, G) = \begin{cases} \min\limits_{i \in C_j(G), i \neq j} \{r_G(i,j)^{-1}\}, & \text{if} \quad |C_j(G)| > 1, \\ 0, & \text{otherwise .} \end{cases} \tag{2.6}$$

Here, $r_G(i, j)$ is the resistance distance between arm i and j in G as in Definition 2.4.1.

Definition 2.4.3 (Competitive and Non-Competitive Arms)**.** *Fix $\boldsymbol{\mu} \in \mathbb{R}^n$, graph D, regularization parameter ρ, confidence parameter δ, and smoothness parameter ϵ. We define $\mathcal{H}_D$ to be the set of competitive arms and $\mathcal{N}_D$ to be the set of non-competitive arms as follows:*

$$\mathcal{H}_D = \left\{ j \in [n] \middle| \Delta_i \leq 2\sqrt{\frac{2}{\rho \mathfrak{I}(i, D)}} \left(2\sigma \sqrt{14 \log \left(\frac{2a_0 n \rho^2 \mathfrak{I}(i, D)^2}{\delta} \right)} + \rho\epsilon \right) \right\} \tag{2.7}$$

and $\mathcal{N}_D \triangleq [n] \setminus \mathcal{H}_D$.

As the name suggests, the arms in $\mathcal{H}$ are close to the optimal arm a^* in mean (competitive compared to the optimal arm a^*) and require several plays before they can be discarded, as shown in the theorem below. Note from the above definition that an arm is more likely to be part of this set if its mean is high (i.e., Δ_i is low) and its influence factor is low. Similarly, the non-competitive set is composed of arms whose means are not competitive with the optimal arm.

Armed with these definitions, we are now ready to state our main theorem that characterizes the performance of GRUB.

2.4.2 Sampling Policy Performance

Cyclic sampling policies have been traditionally used in multi-armed bandit problems for best-arm identification Even-Dar *et al.* (2006). The sample complexity bound for GRUB with cyclic sampling is as follows:

Theorem 2.4.4 (GRUB Sample Complexity)**.** *Consider n-armed bandit problem with mean vector* $\boldsymbol{\mu} \in \mathbb{R}^n$*. Let* $G = (V, E)$ *be the similarity graph with the vertex set* $V = [n]$ *and edge set* E*, let* $\mathcal{G}$ *be the set of subgraphs of* G *, and further suppose that* $\boldsymbol{\mu}$ *is* ϵ*-smooth i.e.,* $\|\boldsymbol{\mu}\|_G \leq \epsilon$*. Define*

$$T_{\textit{sufficient}} \triangleq \arg\min_{D \in \mathcal{G}} \sum_{C \in \mathcal{C}_D} \left[\sum_{\substack{i \in C \cap \mathcal{H}_D \\ i \neq 1}} \frac{1}{\Delta_i^2} \left[c_1 \log \frac{c_2}{\delta \Delta_i} + \frac{\rho \epsilon}{2} \right] + \max_{i \in C \cap \mathcal{N}_D} \frac{2}{\Delta_i^2} \left[c_1 \log \frac{c_2}{\delta \Delta_i} + \frac{\rho \epsilon}{2} \right] \right],$$

where $\Delta_i = \mu^* - \mu_i$ *for all suboptimal arms,* $\mathcal{H}_D$ *and* $\mathcal{N}_D$ *are as in Definition 2.4.3,* $\mathcal{C}_D$ *is the set of connected components of a given graph* D *and* c_1, c_2 *are constants independent of system parameters. Then, with probability at least* $1 - \delta$*, GRUB: (a) terminates in no more than* $T_{\textit{sufficient}}$ *rounds, and (b) returns the best arm* $a^* = \arg\max_i \mu_i$.

Remark 2.4.5. *The required number of samples for the successful elimination of suboptimal arms, and therefore the successful identification of the best arm can be split into two categories based on the sets defined in Definition 2.4.3. Each sub-optimal* highly competitive arm $j \in \mathcal{H}$ *requires* $\mathcal{O}(1/\Delta_j^2)$ *samples, which is comparable to the classical (graph-free) best-arm identification problem. Additionally, the non-competitive arms* $\mathcal{N}$ *can be eliminated without being played, depending on the influence factor: one round of the cyclic sampling suffices to eliminate these arms (even if they are never played!). We refer the reader to Appendix A.4 for a more detailed discussion. Indeed, the smaller* $|\mathcal{H}|$ *is, the more the graph side information benefits GRUB and vice-versa.*

Remark 2.4.6. *Note that* $T_{sufficient}$ *in Theorem 2.4.4 involves the minimum over all subgraphs. As we show in Lemma A.7.8 in the appendix,* $\mathfrak{I}$ *can actually increase if one restricts their attention to certain subgraphs of* G*; this, in turn, increases the size of* $\mathcal{N}$ *and decreases the size of* $\mathcal{H}$*, hence, giving a tighter upper bound on the performance of the algorithm. GRUB automatically adapts to the best subgraph to maximize the influence factor* $\mathfrak{I}(\cdot,\cdot)$ *to obtain the best possible sample complexity and this is reflected in the statement of Theorem 2.4.4.*

The complete proof of Theorem 2.4.4 can be found in Appendix A.4, where we also provide more insights on the behavior of the confidence bound as a function of the number of samples acquired. These results may be of independent interest to the reader.

2.4.3 Improved Sampling Policies

As can be inferred from the pseudocode of Algorithm 1 the primary goal of the `Sampling-Policy` is the quick and safe elimination of suboptimal arms, achieved through shrinking of the confidence bounds $\beta_i(t)\sqrt{(t_{\text{eff},i})^{-1}}$ for all arms i still in

consideration at time t.

Theorem 2.4.4 established guarantees on $T_{\text{sufficient}}$ for naive cyclic sampling policy, i.e. a sampling policy that doesn't directly exploit the graph properties in this arm choice. Note that, even if the sampling policy doesn't utilize any graph properties, the similarity graph is still being utilized in computing the mean estimate and the confidence widths. To enhance the involvement of graph structural information in arm sampling policy, a few alternatives can be characterized:

- **Marginal variance minimization (MVM):** Pick the arm which has the highest confidence bound width. Specifically, at time t, let $\pi_T = \arg\min_{i \in A} t_{\text{eff},i} = \arg\max_{i \in A} [V_T^{-1}]_{ii}$, where A is the set of indices of the arms under consideration.

- **Joint variance minimization – nuclear (JVM-N):** This variant is inspired from the concept of V-optimality Ji and Han (2012). JVM-N picks the arms which leads to a maximum decrease in the value of confidence widths across all arms, in the sense of nuclear norm. Specifically, $\pi_T = \arg\min_{i \in A} \|(V_T + \mathbf{e}_i \mathbf{e}_i^T)^{-1}\|_* - \|V_T^{-1}\|_*$, where $\|\cdot\|_*$ denotes the nuclear norm.

- **Joint variance minimization – operator (JVM-O).** Taking inspiration from Σ-optimality Ma *et al.* (2015b, 2013b), JVM-O picks arms which leads to maximum decrease in the value of confidence widths across all arms in the sense of operator norm. $\pi_T = \arg\min_{i \in A} \|(V_T + \mathbf{e}_i \mathbf{e}_i^T)^{-1}\|_{\text{op}} - \|V_T^{-1}\|_{\text{op}}$

Comparison of the performance of MVM, JVM-N and JVM-O with the baseline of cyclic sampling is provided through synthetic experiments in Section 5.6. In the next section, we derive fundamental lower bounds on the sample complexity *any* algorithm requires in order to solve the said problem

2.5 Lower Bounds

Let us consider an n-armed bandit setup with arm indices $[1, \ldots, n]$. Let μ^* indicate the mean of the optimal arm and μ_i indicate the mean values of all other arms such that $\mu_i < \mu^*$. For the rest of this section, without loss of generality, let the index of the optimal arm be 1.

Theorem 2.5.1. *Given an n-armed bandit model with associated mean vector $\boldsymbol{\mu} \in \mathbb{R}^n$ and similarity graph G smooth on $\boldsymbol{\mu}$, i.e. $\langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle \leq \epsilon$, for any $0 < \epsilon < \epsilon_0$. Let $G = ([n], E)$ be the graph with only isolated cliques and w.l.o.g let arm 1 be the optimal arm. Then define*

$$T_{necessary} = \sum_{C \in \mathcal{C}_G / C^*} \min_{j \in C} \left\{ \frac{4\sigma^2 \log 5}{(\Delta_j - \sqrt{\epsilon})^2} \right\} + \sum_{j \in C^*/1} \frac{4\sigma^2 \log 5}{\Delta_j^2}, \tag{2.8}$$

where C^* is the clique with the optimal arm and $\epsilon_0 := \min_{i \in [n]/1, j \in C(i)} \left[\Delta_j \left[1 - \frac{\Delta_i}{\sqrt{\Delta_i^2 + \Delta_j^2}} \right] \right]^2$. Then any δ-PAC algorithm will need at-least $T_{necessary}$ steps to terminate, provided $\delta \leq 0.1$.

Using Theorem 2.5.1, we can show that GRUB is minimax optimal for a n-armed bandit problems for a certain class of similarity graph G. The following result shows that the upper bound on the sample complexity provided in Theorem 2.4.4 matches the lower bound established in Theorem 2.5.1 in Δ_i up to a constant factor.

Corollary 2.5.2 (Isolated clusters)**.** *Consider the setup as in Theorem 2.5.1 with the further restriction that graph G be such that the optimal node is isolated and $\epsilon < \min_{j \in [n]} \frac{\Delta_j^2}{2}$. Define,*

$$T_{necessary} \geq \sum_{C \in \mathcal{C}_G / \{1\}} \max_{j \in C} \left\{ \frac{8\sigma^2 \log 5}{\Delta_j^2} \right\}. \tag{2.9}$$

Then any algorithm that takes fewer than $T_{necessary}$ samples will have a probability of error at least 0.1.

As can be seen in Corollary 2.5.2, the lower bound expression can scale as standard n-armed bandit (implying no added advantage of having graph side-information) or can behave as a $|\mathcal{C}_G|$-armed bandit problem (scales as the number of clusters in graph G rather than number of nodes n) purely by changing the similarity graph G. The difference between $\mathcal{C}_F$ (connected components in the subgraph constructed by making optimal arm isolated) and $\mathcal{C}_G$ (connected components in the given similarity graph) can lead to more interesting behavior in terms of lower bound expressions on sample complexity.

2.6 ζ Best Arm Identification

It can be observed from Theorem 2.4.4 that the fact that the means are ϵ-smooth implies that distinguishing arm j from a^* would require at least $O(\epsilon^{-2})$ samples. A tighter upper bound on the violation ϵ and an edge between j and a^* would make the suboptimal arm j harder to eliminate. However, it stands to reason that in such situations, it might be more practical to not demand for the absolute best arm, but rather an arm that is nearly optimal. Indeed, in several modern applications, we discuss in Section **??**, finding an approximate best arm is tantamount to solving the problem. In such cases, a simple modification of GRUB can be used to quickly eliminate definitely suboptimal arms, and then output an arm that is guaranteed to be nearly optimal. To formalize this, we consider the ζ-best arm identification problem as follows.

Definition 2.6.1. *For a given $\zeta > 0$, arm i is called ζ-best arm if $\mu_i \geq \mu_{a^*} - \zeta$, where $a^* = \arg\max_i \mu_i$*

The goal of the ζ-best arm identification problem is to return an arm $\tilde{a}$ that is $\zeta-$optimal. We achieve this by a simple modification to GRUB, which we dub

$\zeta-$GRUB, which ensures that all the remaining arms i satisfy $4\beta(t_i)\sqrt{t_{\text{eff},i}^{-1}} \leq \zeta$. It then outputs the best arm amongst those that are remaining. The following theorem characterizes the sample complexity for ζ-GRUB:

Theorem 2.6.2. *Consider n-armed bandit problem with mean vector $\boldsymbol{\mu} \in \mathbb{R}^n$. Let G be the given similarity graph on the vertex set $[n]$, and further suppose that $\boldsymbol{\mu}$ is ϵ-smooth. Let $\mathcal{C}$ be the set of connected components of G. Define,*

$$T_{\textit{sufficient}} \triangleq \arg\min_{D \in \mathcal{G}} \sum_{C \in \mathcal{C}_D} \left[\sum_{i \in C \cap \mathcal{H}_D} \frac{1}{(\Delta_i \vee \zeta)^2} \left[c_1 \log \frac{c_2}{\delta(\Delta_i \vee \zeta)} + \frac{\rho\epsilon}{2} \right] + \max_{i \in C \cap \mathcal{N}_D} \left\{ \frac{2}{(\Delta_i \vee \zeta)^2} \left[c_1 \log \frac{c_2}{\delta(\Delta_i \vee \zeta)} + \frac{\rho\epsilon}{2} \right] \right\} \right], \tag{2.10}$$

where $\Delta_i = \mu^* - \mu_i$ for all suboptimal arms, $\mathcal{H}_D$ and $\mathcal{N}_D$ are as in Definition 2.4.3, $\mathcal{C}_D$ is the set of connected components of a given graph $D$and $\Delta_i \vee \zeta = \max\{\zeta, \Delta_i\}$ and c_1, c_2 are constants independent of system parameters. Then, with probability at least $1-\delta$, ζ-GRUB: (a) terminates in no more than $T_{\textit{sufficient}}$ rounds, and (b) returns a ζ-best arm.

The pseudocode for the ζ-GRUB is as below :

Algorithm 2 ζ-GRUB

Input: Regularization parameter ρ, Smoothness parameter ϵ, Error bound δ, Total arms n, Laplacian L_G, Sub-gaussianity parameter σ

$t \leftarrow 0$

$A = \{1, 2, \ldots, n\}$

$t = 0$

$V_0 \leftarrow \rho L_G$

$\mathcal{C}(G) \leftarrow$ `Cluster-Identification`(L_G)

for $C \in \mathcal{C}(G)$ **do**

  $t \leftarrow t + 1$

  Pick random arm $k \in C$ to observe reward $r_{t,k}$

  $V_t \leftarrow V_{t-1} + \mathbf{e}_k \mathbf{e}_k^T$, and $\mathbf{x}_t \leftarrow \mathbf{x}_{t-1} + r_{t,k}\mathbf{e}_k$

end for

while $|A| > 1$ **do**

  $t \leftarrow t + 1$

  $\beta(t) \leftarrow 2\sigma\sqrt{14 \log\left(\frac{2n(t+1)^2}{\delta}\right)} + \rho\epsilon$

  $k \leftarrow$ `Sampling-Policy`$(t, V_t, A, \mathcal{C}(G))$

  Sample arm k to observe reward $r_{t,k}$

  $V_t \leftarrow V_{t-1} + \mathbf{e}_k \mathbf{e}_k^T$

  $\mathbf{x}_t \leftarrow \mathbf{x}_{t-1} + r_{t,k}\mathbf{e}_k$

  $\hat{\boldsymbol{\mu}}_t \leftarrow V_t^{-1}\mathbf{x}_t$

  $a_{\max} \leftarrow \arg\max_{i \in A} \left[\hat{\mu}_t^i - \beta(t_i)\sqrt{[V_t^{-1}]_{ii}}\right]$

  $A \leftarrow \Big\{\mathbf{a} \in A \mid \hat{\mu}_t^{a_{\max}} - \hat{\mu}_t^a \leq \beta(t_a)\sqrt{[V_t^{-1}]_{aa}}$

  $\qquad + \beta(t_{a_{\max}})\sqrt{[V_t^{-1}]_{a_{\max}a_{\max}}}\Big\}$

  $A \leftarrow A / \left\{a \in A \mid \beta(t_a)\sqrt{[V_t^{-1}]_{aa}} \leq \frac{\zeta}{2}\right\}$

end while

return $\arg\max \left\{\mu_i |\ i \in \{a \in [n] | \beta(t_a)\sqrt{[V_t^{-1}]_{aa}} \leq \frac{\zeta}{2}\} \cup A\right\}$;

2.7 Experiments

For all our experiments, we use a standard laptop with Intel® Core™ i7-10875H CPU @ 2.30GHz × 16 with 32 GB memory. We set the probability of error $\delta = 1e-3$, the penalizing constant $\rho = 2.0$, and noise variance of the subgaussian distribution $\sigma = 2.0$..For the additional graph information, we consider 2 cases: G is a Stochastic Block model(SBM) with parameters $(p, q) = (0.9, 1e^{-4})$ and G is a Barabási–Albert(BA) graph with parameter $m = 2$, both containing 10 clusters. We record the stopping time for 20 runs and plot the results. We evaluate GRUB with different sampling strategies from section 2.4.3 and compare its performance to standard UCB algorithm (Lattimore and Szepesvári, 2020). The full code used for conducting experiments can be found at the following Github repository.

Figure 2.1 compares the baseline cyclic algorithm (UCB algorithm without graph information) with GRUB and its variants (GRUB-MVM, JVM-O, JVM-N) as listed in Section 2.4.3. The x-axis represents the number of arms while keeping the number of clusters constant. As can be seen, all the graph-based methods keep performing better compared to standard baseline UCB, which shows almost linear growth with the number of arms. Interestingly, note that JVM-O, JVM-N, and MVM perform better with an increase in the number of arms in the bandit problem. This is attributed to the fact that increasing the number of arms while keeping the number of clusters static increases the *density* of connections per arm and thereby improving performance.

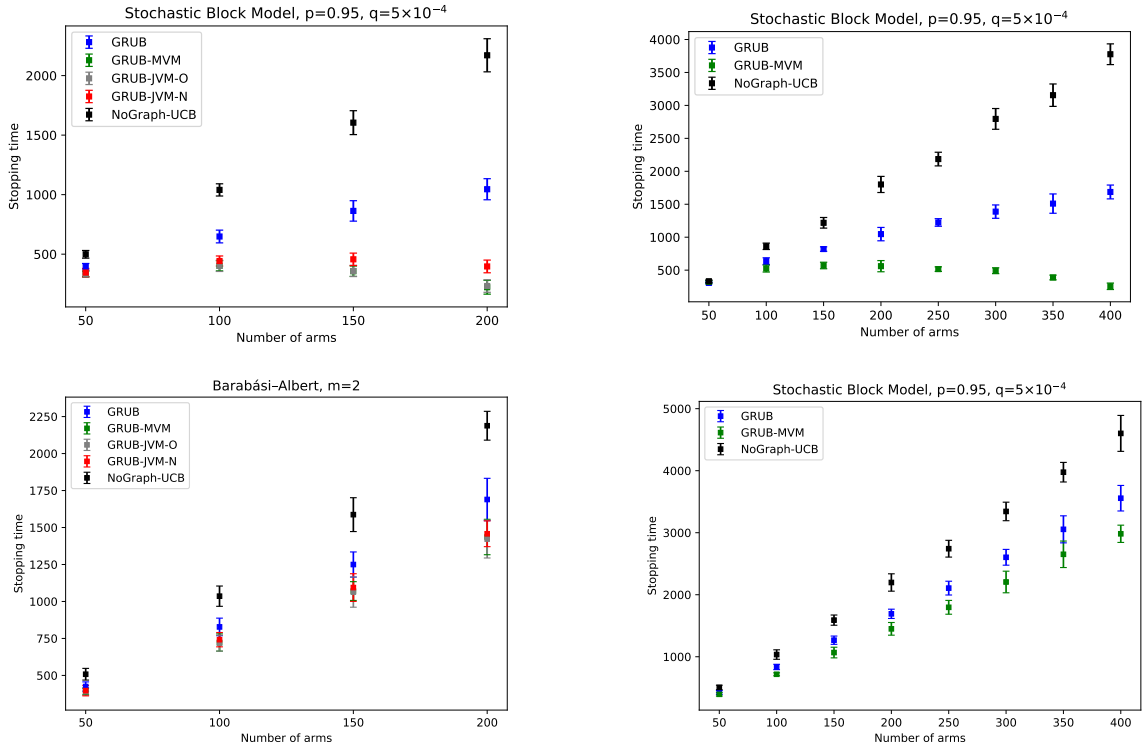


Figure 2.1: (Best seen in color) Stopping time vs number of arms of GRUB using various sampling protocols for SBM ($(p, q) = (0.95, 1e-4)$) [Top] and BA ($m = 2$) [Bottom] . Graph based pure exploration methods outperform the standard cyclic UCB method in terms of stopping time

Real Dataset: It is difficult to obtain a published dataset that exactly fits our problem of pure exploration with graph side information. In order to create a semi-real problem setup, we append an already existing network of users with a corresponding (synthetic) mean structure so as to satisfy the graph side information constraint. We use graphs from SNAP Leskovec and Krevl (2014) for these experiments. We sub-sample the graphs using Breadth-First Search (to retain connected components) to generate the graphs for our experiments. We use the LastFM Rozemberczki and Sarkar (2020), subsampled to 229 nodes, and Github Social Rozemberczki *et al.* (2019) subsampled to 242 nodes.

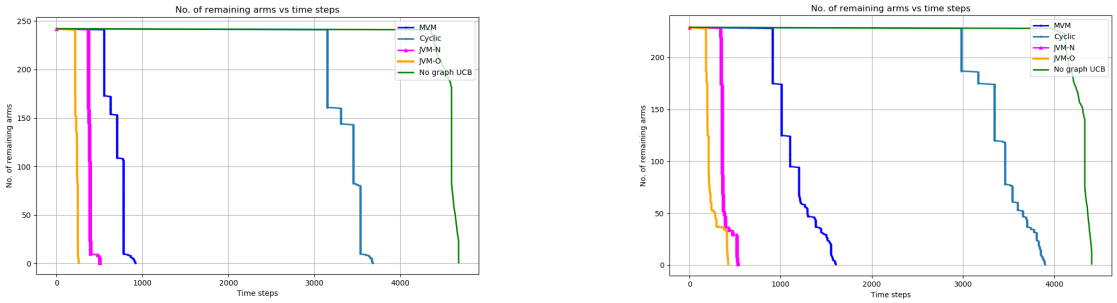


Figure 2.2: (Best seen in color) Cardinality of $|A_t|$ vs. time t of GRUB using different sampling protocols for Github social graph (left) and LastFM graph (right). With no graph information, UCB requires orders of magnitude more samples compared to policies that use explicitly graph information. The cyclic sampling policy is not as competitive on real world datasets

Figure 2.2 plots the number of arms still in consideration $|A_t|$ vs. time t for a single run of the pure exploration problems. This provides us better insights into the behavior of GRUB with different sampling protocols (Section 2.4.3) and standard UCB approach. In all the experiments, it is evident that GRUB with any of the sampling policies outperform UCB algorithm Lattimore and Szepesvári (2020), which does not leverage the graph. Further, within the various sampling policies, the MVM sampling policy seems to outperform other sampling policies (Figure 2.2). For both Github and LastFM datasets, the MVM policy obtains the best arm in ~ 300 rounds compared to traditional UCB that takes ~ 4500 rounds. A rigorous theoretical characterization of the above sampling policies is an exciting avenue for future research.

2.8 Comparison to Linear Bandits

In this section, we provide a toy example as well as a theoretical base to show the difference between the framework of bandits with graph side information and linear

bandits. In this appendix, we first explain the working of the toy example in more detail and then head towards the proof of proposition.

2.8.1 Toy Example

Consider 3-armed bandit problem with graph side information: Let graph G encode the similarity relation between the mean values of the three arms, i.e.

$$\langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle \leq \epsilon \tag{2.11}$$

for some constant $\epsilon > 0$. Let E_G denote the edge set of graph G and $\mathbb{1}_{(1,2)}, \mathbb{1}_{(2,3)}$ and $\mathbb{1}_{(1,3)}$ encodes the event if edges $\{(1,2),(2.3),(1,3)\} \in E_G$ are present in graph G. For the sake of a non-trivial analysis, we take that either $(1,3)$ or $(2,3)$ is present in E_G (alternate case is argued later).

We can write equation (2.11) as,

$$\mathbb{1}_{(1,2)}(\mu_1 - \mu_2)^2 + \mathbb{1}_{(2,3)}(\mu_2 - \mu_3)^2 + \mathbb{1}_{(1,3)}(\mu_1 - \mu_3)^2 \leq \epsilon \tag{2.12}$$

In order to compare the dependence behaviour of μ_3 on μ_1, μ_2 we can rearrange the above as,

$$\begin{aligned} &\mu_3^2 \left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right) - 2\mu_3 \left(\mathbb{1}_{(2,3)}\mu_2 + \mathbb{1}_{(1,3)}\mu_1\right) \\ &+ \left(\mathbb{1}_{(2,3)}\mu_2^2 + \mathbb{1}_{(1,3)}\mu_1^2 + \mathbb{1}_a(\mu_1 - \mu_2)^2 - \epsilon\right) \leq 0 \end{aligned} \tag{2.13}$$

Looking at equation (2.13) as a quadratic in μ_3 and finding the solutions, we obtain

that,

$$\mu_3 \geq \frac{\left(\mathbb{1}_{(2,3)}\mu_2 + \mathbb{1}_{(1,3)}\mu_1\right)}{\left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right)} - \frac{\sqrt{\left(\mathbb{1}_{(2,3)}\mu_2 + \mathbb{1}_{(1,3)}\mu_1\right)^2 - \left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right)\left(\mathbb{1}_{(2,3)}\mu_2^2 + \mathbb{1}_{(1,3)}\mu_1^2 + \mathbb{1}_{(1,2)}(\mu_1 - \mu_2)^2 - \epsilon\right)}}{\left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right)}$$

$$\mu_3 \leq \frac{\left(\mathbb{1}_{(2,3)}\mu_2 + \mathbb{1}_{(1,3)}\mu_1\right)}{\left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right)} + \frac{\sqrt{\left(\mathbb{1}_{(2,3)}\mu_2 + \mathbb{1}_{(1,3)}\mu_1\right)^2 - \left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right)\left(\mathbb{1}_{(2,3)}\mu_2^2 + \mathbb{1}_{(1,3)}\mu_1^2 + \mathbb{1}_{(1,2)}(\mu_1 - \mu_2)^2 - \epsilon\right)}}{\left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right)} \tag{2.14}$$

Further simplifying it, we get the following:

$$\mu_3 \geq \frac{\left(\mathbb{1}_{(2,3)}\mu_2 + \mathbb{1}_{(1,3)}\mu_1\right)}{\left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right)} - \frac{\sqrt{\epsilon\left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right) - \mathbb{1}_{(2,3)}\mathbb{1}_{(1,3)}(\mu_2 - \mu_1)^2 - \mathbb{1}_{(1,2)}\left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right)(\mu_1 - \mu_2)^2}}{\left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right)}$$

$$\mu_3 \geq \frac{\left(\mathbb{1}_{(2,3)}\mu_2 + \mathbb{1}_{(1,3)}\mu_1\right)}{\left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right)} + \frac{\sqrt{\epsilon\left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right) - \mathbb{1}_{(2,3)}\mathbb{1}_{(1,3)}(\mu_2 - \mu_1)^2 - \mathbb{1}_{(1,2)}\left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right)(\mu_1 - \mu_2)^2}}{\left(\mathbb{1}_{(2,3)} + \mathbb{1}_{(1,3)}\right)} \tag{2.15}$$

The above equation leads to non-trivial bounds as ϵ satisfies equation (2.12).

For the simple case of when edge $(2,3)$ is present and $(1,3)$ is not, the above analysis simplifies to,

$$\mu_2 - \sqrt{\epsilon - \mathbb{1}_{(1,2)}(\mu_1 - \mu_2)^2} \leq \mu_3 \leq \mu_2 + \sqrt{\epsilon - \mathbb{1}_{(1,2)}(\mu_1 - \mu_2)^2} \tag{2.16}$$

A similar analysis can be made when edge $(1,3)$ is present and $(2,3)$ is not.

This shows that even with the complete knowledge of μ_1, μ_2 we can only estimate μ_3 to a interval $[\mu_{\text{low}}, \mu_{\text{high}}]$ where the endpoints of interval are given by equation (2.15).

For the case when neither of the edges $(1,3)$ or $(2,3)$ are present (i.e. $\mathbb{1}_{b,c} = 0$), then $\mu_3 \in [-\infty, \infty]$ even with the full knowledge of μ_1, μ_2 as there is no relation between the means of arm 3 to arm $1, 2$, this is also reflected in the equation (2.15)
We first formally rewrite the two setups:

Bandits with graph side information

Consider an n-armed linear bandit problem, each arm $i \in [n]$ is associated with a mean vector $\mu \in \mathbb{R}^n$, where μ_i corresponds to the mean value of arm i. We are provided with further information using a graph G that $\langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle \leq \epsilon$, where $\epsilon > 0$. In each round t, the learner chooses some arm $i \in [n]$ and observes the reward $y_t = \mu_i + \eta_t$, where η_t is a subgaussian random noise with σ^2 variance. Denote the arm with the best mean reward with i^*, i.e. $i^* = \arg\max_{i \in [n]} \mu_i$. The goal of the learner is to output the index of the arm i^* with probability $1 - \delta$, $\delta > 0$ in as few samples as possible.

Linear bandits

Consider an n-armed linear bandit problem, each arm $i \in [n]$ is associated with a feature vector $\mathbf{x}_i \in \mathbb{R}^d$, where d can be lower than n. In each round t, the learner chooses an action $\mathbf{a}_t = \mathbf{x}_i$ for some $i \in [n]$ and observes the reward $y_t = \langle \mathbf{a}_t, \boldsymbol{\theta} \rangle + \eta_t$, where $\boldsymbol{\theta} \in \mathbb{R}^d$ is an unknown parameter and the η_t is a subgaussian random noise with σ^2 variance. Denote the arm with the best mean reward with i^*, i.e. $i^* = \arg\max_{i \in [n]} \langle \mathbf{x}_i, \boldsymbol{\theta} \rangle$. The goal of the learner is to output the index of the arm i^* with probability $1 - \delta$, $\delta > 0$ in as few samples as possible.

Graph vs Linear Bandit framework

In this section, we address the question of whether the n armed bandit problem, with the additional information of $\langle \boldsymbol{\mu}, L \boldsymbol{\mu} \rangle \leq \epsilon$ can be solved using a linear bandits

framework. The metric we use for such a comparison is the set of n-armed bandit problems i.e. set of $\boldsymbol{\mu}$ which can be expressed once the parameters of the two frameworks are fixed. For the case of linear bandits this would be the lower dimension k and feature vector $\mathbf{a}$ corresponding to the reward and for graph bandit framework this indicates the graph G and ϵ. Let the set of problems addressed by the linear bandit framework be denoted by $L_{\mathbf{a}}$ and that by the graph bandits framework denoted by $L_{G,\epsilon}$. We prove that $L_{\mathbf{a}}$ and $L_{G,\epsilon}$ represent sets with fundamentally different properties. Hence we prove that the set of problems addressed by linear bandits and the proposed graph bandit framework of this paper are fundamentally different as there cannot exist one-to-one mapping between the two.

We can further provide additional arguments for the case when $\langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle = 0$. For this, we demonstrate an example graph bandit problem that is cast as a linear bandit to reveal the incomparability of these frameworks.

Firstly, a n-armed bandit problem without any graph can be easily seen as linear bandits by associating the canonical basis for $\mathbb{R}^n$ $\{\mathbf{e}_i\}_{i=1}^n$ as the feature vectors and the mean vector $\boldsymbol{\mu} \in \mathbb{R}^n$ as the unknown reward vector. This provides up with the mean reward function for arm $i \in [n]$ as $\langle \mathbf{e}_i, \boldsymbol{\mu} \rangle = \mu_i$.

In order to cast the graph bandit problem in a linear bandit framework, we need to associate every arm index i with a feature vector $\mathbf{x}_i$ and identify the unknown feature vector $\boldsymbol{\theta}$ for the problem. We achieve this by modifying the feature vectors $\{\mathbf{e}_i\}_{i=1}^n$ and the reward vector $\boldsymbol{\mu}$ based on the graph Laplacian L_G.

Following is the information available at hand in the current graph bandit problem: we are provided with an n-armed bandit with an unknown mean vector $\boldsymbol{\mu}$ smooth on a graph G, i.e. $\langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle \leq \epsilon$. For this toy problem, we consider the graph G to be connected.

Let $\{\boldsymbol{\nu}_i\}_{i=1}^n$ and $0 = \lambda_1 < \cdots < \lambda_n$ denote the eigenvectors and eigenvalues of

the Laplacian L_G respectively. It can be easily seen that $\boldsymbol{\mu} = \sum_{i=1}^{n} a_i \boldsymbol{\nu}_i$ for some $a_i \geq 0 \;\; \forall i \in [n]$. The reward function of arm j is

$$\langle \mathbf{e}_j, \boldsymbol{\mu} \rangle = \sum_{i=1}^{n} a_i \langle \mathbf{e}_j, \boldsymbol{\nu}_i \rangle = a_1 + \sum_{i=2}^{n} a_i \langle \mathbf{e}_j, \boldsymbol{\nu}_i \rangle$$

the second equality follows from the properties of graph Laplacian we know that $\boldsymbol{\nu}_1 = \mathbb{1}_n$, is the only eigenvector associated with 0 eigenvalues in a connected graph.

Without loss of generality, we can assume $a_1 = 0$ as a_1 does not depend on the arm index j. Notice that letting $a_1 = 0$ is equivalent to having $\sum_{i=1}^{n} \mu_i = 0$. Also, the graph constraint can be rewritten as follows:

$$\langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle \leq \epsilon \Rightarrow \sum_{i=1}^{n} \lambda_i a_i^2 = \langle \boldsymbol{\theta}, \boldsymbol{\theta} \rangle = \|\boldsymbol{\theta}\|_2^2 \leq \epsilon$$

where $\boldsymbol{\theta} = (\sqrt{\lambda_1} a_1, \ldots, \sqrt{\lambda_n} a_n)$.

Using the above we can cast the graph bandit problem as the linear bandit problem with the mean reward function of arm j expressed as

$$\langle \mathbf{e}_j, \boldsymbol{\mu} \rangle = \sum_{i=2}^{n} \frac{\theta_i}{\sqrt{\lambda_i}} \langle \mathbf{e}_j, \nu_i \rangle = \langle \mathbf{x}_j, \boldsymbol{\theta} \rangle$$

Hence, the new linear bandit problem is such that the set of arms is $\{\mathbf{x}_j\}_{j=1}^{n}$, the unknown parameter is a vector $\boldsymbol{\theta}$, the expected reward of an arm is $\langle \mathbf{x}_j, \boldsymbol{\theta} \rangle$ and the unknown parameter satisfies the constraint $\|\boldsymbol{\theta}\|_2^2 \leq \epsilon$.

We discuss below the drawbacks of casting a graph bandit problem into a linear bandit framework:

- The original best-arm identification is an n-armed problem and the recasted linear bandit problem still has feature vectors with dimensionality n hence no low-dimensional benefit of linear bandits is completely lost. Having a performance bound for any algorithm for linear bandits that scales in n, the number of arms gives us no additional advantage.

- The above conversion to linear bandit setup only works when the graph G is connected. Recasting problem setup with disconnected components requires an assumption of $\sum_{i \in C} \mu_i = 0$ on individual connected components, which is unrealistic. The results of GRUB hold with or without this assumption.

- Consider the corner case of $\epsilon = 0$, the linear bandit problem setup derived becomes that of $\arg\max_i \langle \mathbf{x}_i, \boldsymbol{\theta} \rangle$ such that $\|\boldsymbol{\theta}\| \leq 0$ which is only possible if $\|\boldsymbol{\theta}\| = 0$ and in this case we can observe two interesting facts:

  - If the graph G is completely connected then the problem is trivial, since

    $$\epsilon = 0 \Rightarrow \langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle = 0 \Rightarrow (\mu_i - \mu_j)^2 = 0 \quad \forall i, j \in [n], i \neq j$$

    This implies all arms are equal and optimal and the solution is trivial. Here the mean reward function of all arms i is $\langle \mathbf{x}_i, \boldsymbol{\theta} \rangle = 0$ since $\theta = 0$ and hence gives the correct output (any arm i).

  - Suppose graph G has two connected components C_1, C_2, where C_k indicates the arm indices in the connected component k. Further assume that $\mu_i = 1 \quad \forall i \in C_1, \mu_i = -1 \quad \forall i \in C_2$. Considering the case of $\epsilon = 0$ here gives us the following :

    $$\epsilon = 0 \Rightarrow \langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle = 0 \Rightarrow (\mu_i - \mu_j)^2 = 0 \quad \forall i \neq j, i, j \in C_k, k = 1, 2$$

    Here the mean reward function of all arms i is $\langle \mathbf{x}_i, \boldsymbol{\theta} \rangle = 0$ since $\theta = 0$ but this is incorrect since not all arms are optimal.

  Our graph bandit setup and the performance of GRUB is independent of all of these drawbacks and provides us with a better sample complexity than vanilla best arm identification algorithms.

We solidify these arguments with the following Propositions.

Proposition 2.8.1. *Consider n-armed bandit setup parameterized by mean vector* $\boldsymbol{\mu} \in \mathbb{R}^n$. *Given graph G and* $\epsilon > 0$, *let* $D_{G,\epsilon} := \{\boldsymbol{\mu} \in \mathbb{R}^n | \ \langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle \leq \epsilon\}$ *represent a subset of bandit problems in* $\mathbb{R}^n$, L_G *denoting the laplacian matrix corresponding to graph G. Then* $m(D_{G,\epsilon}) > 0$ *where* $m(\cdot)$ *is the Lebesgue measure on* $\mathbb{R}^n$.

Sketch of proof : We solidify the intuition from toy example 2.8.1 to show the distinction in the two frameworks using measure theoretic argument. We split the n-armed bandit problem with graph-side information into two complementary scenarios:

(a) $L_> : \{\boldsymbol{\mu} \in \mathbb{R}^n | \ \ 0 < \langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle \leq \epsilon, \epsilon > 0\}$

(b) $L_= : \{\boldsymbol{\mu} \in \mathbb{R}^n | \ \ \langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle = 0\}$

We prove that the set $D_{G,\epsilon} = L_> \cup L_=$ has fundamentally different measure theoretic properties than $D_{\boldsymbol{\theta}}$ (linear bandits framework) and hence the two problems setups tackle completely different domain of questions.

Theorem 2.8.2. *Consider n-armed bandit setup. Let* $D_{G,\epsilon}, D_{\boldsymbol{\theta}}$ *represent the subset of problems in* $\mathbb{R}^n$ *as follows:*

$$\begin{aligned} D_{G,\epsilon} \ &= \{\boldsymbol{\mu} \in \mathbb{R}^n | \ \ 0 < \langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle \leq \epsilon, \epsilon > 0\} \\ D_{\boldsymbol{\theta}} \ &= \{\boldsymbol{\mu} \in \mathbb{R}^n | \ \ \mu_i = \langle \mathbf{a}_i, \boldsymbol{\theta} \rangle, \quad \mathbf{a}_i \in \mathbb{R}^k, \quad \forall i \in [n]\} \end{aligned}$$

where $k < n, \boldsymbol{\theta} \in \mathbb{R}^k$ *indicates the reward vector and* L_G *is the laplacian matrix corresponding to graph G. Then* $D_{G,\epsilon} \not\subset D_{\boldsymbol{\theta}}$, $D_{\boldsymbol{\theta}} \not\subset D_{G,\epsilon}, m(D_{G,\epsilon}) > 0$ *and* $m(D_{\boldsymbol{\theta}}) = 0$ *where* $m(\cdot)$ *is the Lebesgue measure on* $\mathbb{R}^n$.

Proof. The two problem subset definitions represent the following :

a) $D_{G,\epsilon}$ – Given a graph G and violation parameter ϵ, $D_{G,\epsilon}$ represents mean-reward vectors $\boldsymbol{\mu} \in \mathbb{R}^n$ which satisfy the graph bandit setup.

b) $D_{\boldsymbol{\theta}}$ – Given lower dimension $k < n$ and the corresponding reward vector $\boldsymbol{\theta}$, the set $D_{\boldsymbol{\theta}}$ indicates the set of all the mean rewards $\{\mu_i\}_{i=1}^n$ for n-armed bandit setup such that the mean-reward vector can be represented by k-dimensional feature vectors.

First, consider the following arguments :

- Notice that if $\boldsymbol{\alpha}, \boldsymbol{\beta} \in D_{\boldsymbol{\theta}}$ then $c_1\boldsymbol{\alpha} + c_2\boldsymbol{\beta} \in D_{\boldsymbol{\theta}}$ and $\mathbf{0} \in D_{\boldsymbol{\theta}}$, where $\mathbf{0}$ is the all zero vector in $\mathbb{R}^n$. Hence we can conclude $D_{\boldsymbol{\theta}}$ is a subspace of $\mathbb{R}^n$. Since all the elements of the set $D_{\boldsymbol{\theta}}$ can be as a linear map to a k-dimensional subspace constructed out of $\{\mathbf{a}_i\}_{i=1}^n$, $\mathbf{a}_i \in \mathbb{R}^k$ for all $i \in [n]$, hence $D_{\boldsymbol{\theta}}$ is a k-dimensional subspace of $\mathbb{R}^n$. Accordingly, $m(D_{\boldsymbol{\theta}}) = 0$ where m is the Lebesgue measure on the euclidean space $\mathbb{R}^n$.

- Consider the set $D_{G,\epsilon}$ and $\boldsymbol{\mu}$ such that $\langle \boldsymbol{\mu}, L_G\boldsymbol{\mu} \rangle = 0$ (existence of such a $\boldsymbol{\mu}$ is easy to prove by making $\mu_i = \mu_j$ for every edge in G). Given that $\epsilon > 0$, $\exists \delta > 0$ such that $\forall \boldsymbol{\sigma} \in \mathcal{B}(\mathbf{0}, \delta)$,

$$\begin{aligned}
\langle (\boldsymbol{\mu} + \boldsymbol{\sigma}), L_G(\boldsymbol{\mu} + \boldsymbol{\sigma}) \rangle &= \sum_{\{i,j\} \in E_G} A_{ij}(\mu_i + \sigma_i - \mu_j - \sigma_j)^2 \\
&= \sum_{\{i,j\} \in E_G} A_{ij}(\sigma_i - \sigma_j)^2 \qquad (\langle \boldsymbol{\mu}, L_G\boldsymbol{\mu} \rangle = 0 \Rightarrow \mu_i = \mu_j \ \ \forall (i,j) \in E_G) \\
&\leq \|A_G\|_\infty \sum_{\{i,j\} \in E_G} (\sigma_i - \sigma_j)^2 \\
&\leq 4\|A_G\|_\infty \|\boldsymbol{\sigma}\|_2^2
\end{aligned} \tag{2.17}$$

  Taking $\delta < \frac{\epsilon}{4\|A_G\|_\infty}$ proves that $\forall \boldsymbol{\sigma} \in \mathcal{B}(\mathbf{0}, \delta)$, $\langle (\boldsymbol{\mu} + \boldsymbol{\sigma}), L_G(\boldsymbol{\mu} + \boldsymbol{\sigma}) \rangle \leq \epsilon$. Hence $\mathcal{B}(\mathbf{0}, \delta) \subset D_{G,\epsilon}$ implying $m(D_{G,\epsilon}) > m(\mathcal{B}(\mathbf{0}, \delta)) = \delta^n$.

  Further, consider $\boldsymbol{\theta} \in \mathbb{R}^n$ such that $\langle \boldsymbol{\theta}, L_G\boldsymbol{\theta} \rangle = \epsilon$ for some $\epsilon > 0$ hence $\boldsymbol{\theta} \in D_{G,\epsilon}$. Then it is easy to see that $2\boldsymbol{\theta} \notin D_{G,\epsilon}$ as $\langle 2\boldsymbol{\theta}, L_G(2\boldsymbol{\theta}) \rangle = 4\epsilon > \epsilon$.

We can thus conclude that $D_{\boldsymbol{\theta}}$ is a k-dimensional subspace of $\mathbb{R}^n$ which is a measure zero set and $D_{G,\epsilon}$ is a positive measure set which is not closed under multiplication. Hence we can easily see that $D_{G,\epsilon} \not\subset D_{\boldsymbol{\theta}}$, $D_{\boldsymbol{\theta}} \not\subset D_{G,\epsilon}$. □

Theorem 2.8.3. *Consider n-armed bandit setup. Let $D_{G,\epsilon}, D_{\boldsymbol{\theta}}$ represent the subset of problems in $\mathbb{R}^n$ as follows:*

$$\begin{aligned} D_{G,\epsilon} &= \{\boldsymbol{\mu} \in \mathbb{R}^n | \ \ \langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle \leq \epsilon, \epsilon > 0\} \\ D_{\boldsymbol{\theta}} &= \{\boldsymbol{\mu} \in \mathbb{R}^n | \ \ \mu_i = \langle \mathbf{a}_i, \boldsymbol{\theta} \rangle, \ \ \mathbf{a}_i \in \mathbb{R}^k, \ \ \forall i \in [n]\} \end{aligned}$$

where $k < n, \boldsymbol{\theta} \in \mathbb{R}^k$ indicates the reward vector and L_G is the laplacian matrix corresponding to graph G. Then $D_{G,\epsilon} \not\subset D_{\boldsymbol{\theta}}$, $D_{\boldsymbol{\theta}} \not\subset D_{G,\epsilon}, m(D_{G,\epsilon}) > 0$ and $m(D_{\boldsymbol{\theta}}) = 0$ where $m(\cdot)$ is the Lebesgue measure on $\mathbb{R}^n$.

Proof. We can split the argument into two parts:

- $\langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle > 0$
- $\langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle = 0$

Theorem 2.8.2 addresses the first part of the argument

For the case when second part, i.e. $\langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle = 0$, let $L_= : \{\boldsymbol{\mu} \in \mathbb{R}^n | \ \ \langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle = 0\}$. Note that for any $\boldsymbol{\mu} \in \mathbb{R}^n$ only happens if and only if $\boldsymbol{\mu} \in \mathcal{N}(L_G)$ where $\mathcal{N}(\cdot)$ represents the null space of the matrix. Since L_G is rank deficient $L_=$ is a set in a subspace of $\mathbb{R}^n$ and hence $m(L_=) = 0$.

Thus $m(L_> \cup L_=) > 0$ which is fundamentally different from linear bandits addresses problems of measure zero.

□

Thus we can conclude that the two frameworks of graph and linear bandits address fundamentally different domains of problems.

2.9 Discussion and Broader Impacts

In this work, we consider the problem of best arm identification (and approximate best arm identification) when one has access to information about the similarity

between the arms in the form of a graph. We propose a novel algorithm GRUB for this important family of problems and establish sample complexity guarantees for the same. In particular, our theory explicitly demonstrated the benefit of this side information (in terms of the properties of the graph) in quickly locating the best or approximate best arms. We support these theoretical findings with experimental results in both simulated and real settings.

Future Work and Limitations. We outline several sampling policies inspired by our theory in Section 5.6; an extension of our theoretical results to account for these improved sampling policies is a natural candidate for further exploration. The algorithms and theory of this paper assume knowledge of (an upper bound) on the smoothness of the reward vector with respect to the graph. While this is where one uses domain expertise, this could be hard to estimate in certain real world problems. A generalization of the algorithmic and theoretical framework proposed here that is *adaptive* to the unknown graph-smoothness is an exciting avenue for future work (Cai and Yuan, 2012; Banerjee *et al.*, 2020). The sub-Gaussianity assumption of this work can also be generalized to other tail behaviors in follow up work. Another limitation of this work is that the statistical benefit of the graph-based quadratic penalization comes at a computational cost – each mean estimation step involves the inversion of an $n \times n$ matrix which has a complexity of $O(n^2 \log(n))$. However, an exciting recent line of work suggests that this matrix inversion can be made significantly faster when coupled with a spectral sparsification of the graph G (Vishnoi *et al.*, 2013; Spielman and Teng, 2010) while controlling the statistical impact of such a modification. In the context of this problem, this suggests a compelling avenue for future work that studies the statistics-vs-computation tradeoffs in using graph side information.

For this work, we demonstrated the advantages of this side information in pure exploration problems, given knowledge of such an ϵ. Extensions that consider goodness-

of-fit and misspecification with respect to the graph G and smoothness parameters ϵ are interesting avenues for follow up work. Finally, we focus on the ridge-type regularizer of the form $\langle \mu, L_G \mu \rangle$. For future work, it may be productive to expand to a much broader class of regularizers such as those of the form of $\|A\mu\|_q^p$, where A represents an information/ structural constraint matrix and p, q are some positive numbers.

Potential Negative Social Impacts. Our methods can be used for various applications such as drug discovery, advertising, and recommendation systems. In scientifically and medically critical applications, the design of the reward function becomes vital as this can have a significant impact on the output of the algorithm. One must take appropriate measures to ensure a fair and transparent outcome for various downstream stakeholders. With respect to applications in recommendation and targeted advertising systems, it is becoming increasingly evident that such systems may exacerbate polarization and the creation of filter-bubbles. Especially, the techniques proposed in this paper could reinforce emerging polarization (which would correspond to more clustered graphs and therefore better recommendation performance) when used in such contexts. It will of course be of significant interest to mitigate such adverse outcomes by well-designed interventions or by considering multiple similarity graphs that capture various dimensions of similarity. This is a compelling avenue for future work.

Chapter 3

QUADRATIC FEASIBILITY

This study addresses the quadratic feasibility problem, aiming to retrieve a complex vector $\mathbf{x} \in \mathbb{C}^n$ from m quadratic measurements $\langle A_i\mathbf{x}, \mathbf{x}\rangle_{i=1}^m$. Despite its numerous applications, solving this problem is NP-hard, and achieving identifiability poses challenges. Gaussian random measuring matrices are ubiquitous in existing related literature. In contrast, we focus our analysis on deterministic measurement matrices that can possibly reflect physical setups and structural constraints. The removal of randomness raises the challenge of the problem setup in deriving identifiability and retrieval guarantees. To overcome this, we introduce novel conditions for the design process that ensure identifiability and retrieval through first-order descent methods. Upon satisfying these necessary conditions, we guarantee key characteristics of the optimization landscape which enables gradient algorithms to converge to a globally optimal solution with high probability, regardless of initialization. Optimization landscape simulation provides supplementary intuition supporting the necessity of our proposed conditions and demonstrates the superior performance of our deterministic measurement matrices compared to random Gaussian measurements.

3.1 Introduction

Finding a solution to a system of quadratic equations is an important problem with a wide range of applications. It arises in areas such as power system state estimation Wang *et al.* (2016), phase retrieval Candes *et al.* (2015, 2013); Eldar and Mendelson (2014); Balan and Zou (2014), x-ray crystallography Drenth (2007), the turnpike problem Dakic (2000), and unlabeled distance geometry problemsDuxbury *et al.*

(2016); Huang and Dokmanić (2018) among others. Such problems can be reduced to a *quadratic feasibility problem*, where one is concerned with finding a *feasible* vector $\mathbf{x}$ that conforms to a set of quadratic observations of the form $\{\langle A_i\mathbf{x}, \mathbf{x}\rangle\}_{i=1}^m$ with respect to a set $\{A_i\}_{i=1}^m$ of measurement matrices. Formally, it can be cast as:

$$\text{find } \mathbf{x} \quad \text{such that } \langle A_i\mathbf{x}, \mathbf{x}\rangle = c_i, \quad \forall i = 1, 2, \ldots, m. \tag{P1}$$

The quadratic feasibility problem is an instance of quadratically constrained quadratic programs (QCQPs)Park and Boyd (2017), which has enjoyed a long and rich research history dating back to 1941Dines (1941). Given their broad applicability to critical problems, research in QCQPs continues to be of active interest Park and Boyd (2017); Beck and Pan (2017); Beck (2009); Beck and Eldar (2006). Unfortunately, it is known that solving generic QCQPs is an NP-hard problem Sahni (1974). This combined with the lack of tractable duality properties Pólik and Terlaky (2007) has made it hard to establish a sound theoretical framework for understanding the solutions and computing them. However, an extremely productive line of research has instead considered subclasses of QCQPs that are both practically relevant and can be analyzed. While a significant portion of the research literature focuses on QCQPs characterized with Gaussian random matrices, these don't apply to deterministic situations where the constraints represent physical restrictions of the underlying problem setup. In this paper, we identify the necessary and sufficient conditions required for any quadratic feasibility problem, deterministic or random, to be tractable, by which we mean that the problem setup (P1) can uniquely identify the solution as well as retrieve it. We provide a theoretical and intuitive understanding of the established conditions for the identifiability of the solution, connect the existing literature on gaussian random QCQPs with the specified conditions and further provide a novel extension of the results to the case of noisy measurements, which is widely missing in the current

literature.

Finding a solution to (P1) involves two tasks: i) Ensuring the problem is well-defined, i.e. the intended unknown solution being the only solution solving the system of equations, ii) Developing a trackable approach to recover the unknown solution. In this paper, we solve both these issues for a deterministic system of equations as well as provide guardrails as to when such a system of equations can be tagged as solvable. (more on this later)

We start by analyzing quadratic mapping: $\mathbf{x} \to \{\langle A_i\mathbf{x}, \mathbf{x}\rangle\}_{i=1}^m$, and focus on their ability to generate injective maps up-to a phase factor (note that a quadratic function of any matrix $A \in \mathbb{C}^{n\times n}$, i.e. $\langle A\mathbf{x}, \mathbf{x}\rangle$ is invariant to phase shifts). Following up on injective properties, we establish the distance metric to prove isometric properties of quadratic mapping, thereby ensuring uniqueness of solution for (P1) up to a phase constant. Next, we consider the question of computationally tractable approaches to converge to this unique solution. Unlike most other works in literature, for this work we assume that measurements are corrupted with Gaussian noise. A natural approach to tackle (P1) is to reformulate the problem as a loss function. We prove that the loss function is necessarily nonconvex in nature making it NP-hard. Our main result of the paper revolves around showing that under certain conditions, this loss landscape is well-behaved. While such statements of this nature are already established in the literature, we are the first to prove it in the case of deterministic matrix ensemble. Such nature of the loss landscape enables successful global recovery for any first-order algorithm to solve QFP. As a clear connection to the related literature, we also prove that these required conditions hold with very high probability for random measurements (specifically an ensemble of Gaussian random matrices $\{A_i\}_{i=1}^m$ with $m > O(n)$)

The rest of the paper is organized as follows. Section 3.3 discusses about relevant

works in literature. Section 3.4 highlights the main results of this work. We discuss some related work in Section 3.3. In Section 3.5 we establish and analyze isometry/identifiability properties of the mapping $\{\langle A_i\mathbf{x}, \mathbf{x}\rangle\}_{i=1}^m$ when the measurement matrices are complex Gaussian. Finally, Section 3.6 casts the problem as a quadratic loss minimization problem (suitable for efficient algorithms) and establishes favorable properties of the loss landscape that allow one to find a solution using gradient-based methods with arbitrary initial points.

Before we state the main results of the paper, we introduce some notation that will be used throughout the paper.

3.2 Notation

For any $r \in \mathbb{N}$, we write $[r]$ to denote the set $\{1, 2, \ldots, r\}$. We let $\mathbb{C}^n$ and $\mathbb{R}^n$ denote the n-dimensional complex and real vector spaces, respectively. Unless otherwise stated, bold letters such as $\mathbf{x}$ indicate vectors in $\mathbb{C}^n$; $\mathbf{x}_\mathbb{R}$ and $\mathbf{x}_\mathbb{C}$ denote the real and the imaginary part of the vector $\mathbf{x}$, respectively. We denote complex conjugate of $\mathbf{x}$ by $\bar{\mathbf{x}}$. Capital letters such as X denote matrices in $\mathbb{C}^{n\times n}$. The use of i (without serif) indicates the complex square root of -1 (we will use i to indicate an indexing variable). We let $\mathcal{S}^{a,b}(\mathbb{R}^{n\times n})$ denote the set of all matrices $X \in \mathbb{R}^{n\times n}$ having a non-negative eigenvalues and b negative eigenvalues, where $a + b = n$. The set $\mathbf{H}_n(\mathbb{C})$ denotes the set of all $n \times n$ Hermitian matrices. We write $A^\top$ and $A^\dagger$ to denote, respectively, the transpose and the Hermitian transpose (transpose conjugate) of a matrix A. We use $\langle\cdot,\cdot\rangle$ to denote the inner vector product in the complex space. The symmetric outer product, denoted by $[[\cdot,\cdot]]$, is defined as $[[\mathbf{u}, \mathbf{v}]] = \mathbf{u}\mathbf{v}^\dagger + \mathbf{v}\mathbf{u}^\dagger$. Finally, we will let $\sim$ denote the following equivalence relation on $\mathbb{C}^n$: $\mathbf{x} \sim \mathbf{y}$ if and only if $\mathbf{x} = c\mathbf{y}$ for some $c \in \mathbb{C}$ with $|c| = 1$. We will write $\mathbb{C}^n_\angle \triangleq \mathbb{C}^n/\sim$ to denote the associated quotient space. Given a set of matrices $\mathcal{A} = \{A_i\}_{i=1}^m \subset \mathbb{H}_n(\mathbb{C})$, we will let

$\mathcal{M}_{\mathcal{A}}$ denote the following mapping from $\mathbb{C}^n_{\angle} \to \mathbb{C}^m$:

$$\mathcal{M}_{\mathcal{A}}(\mathbf{x}) = (\langle A_1\mathbf{x}, \mathbf{x}\rangle, \langle A_2\mathbf{x}, \mathbf{x}\rangle, \ldots, \langle A_m\mathbf{x}, \mathbf{x}\rangle). \tag{3.1}$$

While $\mathcal{M}_{\mathcal{A}}$ technically operates on the equivalence classes in $\mathbb{C}^n_{\angle}$, we will abuse the notation slightly and think of $\mathcal{M}_{\mathcal{A}}$ as operating on the elements of $\mathbb{C}^n$. Let $d_{\angle}(\cdot,\cdot)$ denote the following Frobenius distance metric for rank-1 matrices,

$$d_{\angle}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\mathbf{x}^\dagger - \mathbf{y}\mathbf{y}^\dagger\|_F \quad \text{for any } \mathbf{x}, \mathbf{y} \in \mathbb{C}^n. \tag{3.2}$$

We further define cross distance between two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ as follows,

$$d_c(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\mathbf{y}^\perp - \mathbf{y}\mathbf{x}^\perp\|_F \quad \text{for any } \mathbf{x}, \mathbf{y} \in \mathbb{C}^n. \tag{3.3}$$

In order to give an intuitive understanding of the core reasoning for some of the concepts in the work, we use the following toy setup.

3.3 Related Works

QCQPs have enjoyed a lot of attention over the last century. When it comes to provably solving, due to the limitation of the duality properties of QCQPs Barvinok (1995), a significant fraction of research has focused predominantly on heuristic approaches to their solution Konar and Sidiropoulos (2017a,b); Konar (2017). Recently, an ADMM-based method has been proposed inHuang and Sidiropoulos (2016) with an asymptotic convergence result based on the duality properties of QCQPs. Our results in this paper bring new insights to this area by analyzing a subset of QCQPs, namely, the Quadratic feasibility problems.

The Quadratic feasibility problem (P1) arises in many applications, including phase retrieval Candes *et al.* (2015) and power system state estimation Wang *et al.*

(2016). Phase retrieval in and of itself finds applications in a wide variety of fields such as imaging, optics, quantum tomography, and audio signal processing with a wide literature, including Eldar and Mendelson (2014); Candes *et al.* (2015); Tan and Vershynin (2017); Balan and Zou (2014). In Candes *et al.* (2013), an approximate ℓ_1 isometry property was established for the phase retrieval problem, but the bounds therein are not strong enough to provide RIP-like guarantees. In this paper, we improve these bounds to establish isometry results for a large class of problems and provide RIP-type bounds.

The authors in Wang and Xu (2017) provide lower bounds on the minimum number of independent measurements required for a successful recovery for the quadratic feasibility problem. Our high probability bounds are ordered optimal in the number of measurements required for successful recovery. More recently, Huang *et al.* (2019) showed that the quadratic feasibility problem Gradient descent can be solved, with high probability, provided a good initialization is used.The current work takes an alternate path by analyzing the landscape of the associated ℓ_2-loss function. We suspect that results, such as the recovery of solution using gradient descent, are a result of underlying properties of the landscape of the function landscape. In particular, for the ℓ_2-loss function, we prove that all local minima are global and all saddle points are strict. Thus, our results enable gradient-based algorithms with arbitrary initialization to recover the solution for the quadratic feasibility problem.

The study of recovery performance for the quadratic feasibility problem has largely centered around phase retrieval problems. The earliest work in this field can be traced back to Balan *et al.* (2006), who established necessary and sufficient conditions for measurement vectors to result in injective and stable intensity measurements. Sun *et al.* (2018) showed that, when the measurements are i.i.d complex Gaussian and the number of measurements is sufficient, there is a high probability that there are no false

local minimizers and all global minimizers are equal to the target signal. Huang and Xu (2020) provided a sharp estimation error of the model without noise distribution assumptions, given that the measurement vectors are random Gaussian.

In the context of the general quadratic measurements model (QMR), Huang *et al.* (2021) and Wang and Xu (2017) introduced the concept of phase retrieval property and explored its connections to low-rank matrix recovery and nonsingular bilinear form. They derived results on the minimum number of measurements needed for matrix recovery and applied these to phase retrieval problems. Specifically, they showed that a set of $p \times p$ matrices has the phase retrieval property if $n \geq 2p - 1$ in the real case and $n \geq 4p - 4$ in the complex case. The concept of almost everywhere phase retrieval property has also been investigated for a set of matrices.

In our previous work (Thaker *et al.* (2020)), we investigated the quadratic feasibility problem and established conditions for its identifiability. We concluded that, if the matrices $A_1, ..., A_n$ are Hermitian matrices sampled from a complex Gaussian distribution, any first-order algorithm from an arbitrary starting point can converge to a globally optimal solution with a high probability.

In the phase retrieval literature, the problem of designing the measurement matrix received considerably less attention compared to the design of retrieval algorithms. An important desirable property that measurement matrices should satisfy is a unique relationship between the signal and the magnitudes of its projections, up to an inherent phase ambiguity. In many works, particularly in theoretical performance analysis of phase retrieval algorithms Candes *et al.* (2013, 2015); Candes (2008), the matrices are assumed to be *random*, commonly with i.i.d. Gaussian entries. However, in practical applications, the measurement matrix corresponds to a fixed physical setup, so it is typically a deterministic matrix, with possibly structural constraints. For example, in optical imaging, lenses are modeled using dft matrices and optical masks correspond

to diagonal matrices Candes *et al.* (2015). Measurements based on oversampled dft matrices were studied in Huang *et al.* (2016), measurement matrices which correspond to the parallel application of several dfts to modulated versions of the soi were proposed in Candes *et al.* (2015), and Drémeau *et al.* (2015) studied phase recovery using fixed binary measurement matrices, representing hardware limitations in optical imaging systems.

All the works above considered *noiseless observations*, hence, the focus was on obtaining uniqueness of the magnitudes of the projections in order to guarantee recovery, though the recovery method may be intractable Eldar and Mendelson (2014). When noise is present, such uniqueness no longer guarantees recovery, thus a different design criterion should be considered. Recovery algorithms as well as specialized deterministic measurement matrices were considered in several works. In particular, Jaganathan *et al.* (2016); Bendory *et al.* (2017) studied phase recovery from short-time Fourier transform measurements, Pedarsani *et al.* (2017) proposed a recovery algorithm and measurement matrix design based on sparse graph codes for sparse soi taking values on a finite set, Iwen *et al.* (2016) suggested an algorithm using correlation based measurements for flat soi, i.e., strictly non-sparse soi, and Bodmann and Hammen (2014) studied recovery methods and the corresponding measurement matrix design for the noisy phase retrieval setup by representing the projections as complex polynomials.

A feasibility problem is often cast as a minimization problem with a suitably chosen loss function. Even with a nonconvex objective, gradient based methods have proven to work for phase retrieval Tan and Vershynin (2017); Balan and Zou (2014); Candes *et al.* (2015), matrix factorization Bhojanapalli *et al.* (2016a); Jin *et al.* (2016) and robust linear regression Jain *et al.* (2017). The work in Sun *et al.* (2018) has established landscape properties for the phase retrieval problem, which sheds light on the success of gradient based methods in solving the problem. In this work, we extend these

results to a wider class of problems along with additional insights into the problem properties. In Ge *et al.* (2017), it was shown that many nonconvex loss functions have specific landscape properties, which allows gradient based algorithms to recover a globally optimal solution without any additional information. One unfortunately cannot readily transport those results to our setting, mainly due to the significant differences between the real and complex vector spaces. For instance, a quadratic feasibility problem in $\mathbb{R}^n$ has only two isolated local minima, while it has a continuum of minima in $\mathbb{C}^n$.

A natural optimality condition for the noisy setup, without focusing on a specific recovery algorithm, is to design the measurement matrix to minimize the achievable mse in estimating the soi from the observations. However, in phase retrieval, the soi and observations are not jointly Gaussian, which makes computing the mmse for a given measurement matrix in the vector setting very difficult. Furthermore, even in the linear non-Gaussian setting, a closed-form expression for the derivative of the mmse exists only for the scalar case Guo *et al.* (2011), which corresponds to a single observation. Therefore, gradient-based approaches for mmse optimization are difficult to apply as well.

3.4 Main Results

We consider the quadratic feasibility problem (P1) where the aim is to recover a complex decision vector $\mathbf{x} \in \mathbb{C}^n$ through its m noisy quadratic measurements of the form of $\langle A_i\mathbf{x}, \mathbf{x}\rangle$ where $A_i \in \mathbb{C}^{n\times n}$ are Hermitian matrices and the noisy measurements are represented by $c_i \in \mathbb{R}$ for all $i \in [m]$. We start by characterizing conditions under which the quadratic feasibility problem (P1) is well-defined. We define (α, β)-*stability* to establish the uniqueness of the solution $\mathbf{x}^*$ of (P1).

Lemma 3.4.1. *The mapping $\mathcal{M}_{\mathcal{A}}$ is injective iff it is (α, β)-stable for some constants*

$0 < \alpha \leq \beta$.

Having established that (P1) has a uniquely identifiable solution (upto a phase ambiguity), we next turn our attention to finding a feasible solution in a computationally efficient manner, specifically through the usage of local gradient methods (gradient-based methods which do not have access to any global information). To this end, one may consider recasting the quadratic feasibility problem as a quadratic minimization problem of the ℓ_2-loss function, as follows:

$$\min_{\mathbf{x}\in\mathbb{C}^n} f(\mathbf{x}), \qquad f(\mathbf{x}) \triangleq \frac{1}{m}\sum_{i=1}^{m} \left|\langle A_i\mathbf{x},\mathbf{x}\rangle - c_i\right|^2. \tag{F2}$$

Unfortunately, this optimization problem is non-convex (can be seen by analysing the cost function at $\mathbf{x}^*, -\mathbf{x}^*$ and 0) and, in general, one may not guarantee any gradient based method to converge to a global minimum.

We define *κ-cross-stability* in order for local gradient methods to successfully recover a quadratic feasible point. Intuitively κ-cross-stability ensures that the set of measuring matrices $\{A_d\}_1^m$ are capable enough to sense the energy in the direction perpendicular to $\mathbf{x}^*$. We prove that, provided the measuring matrices A_i's satisfy the α,β-stability, and κ-cross-stability, the landscape of said loss function can actually be benign, making it possible for first-order methods to recover the solution.

Theorem 3.4.2 (sketch)**.** *Let $\mathcal{A} = \{A_d\}_1^m$ be the set of measurement matrices for the quadratic feasibility problem (P1) and let them satisfy α,β-stability and κ-cross-stability. Suppose that $\mathbf{z} \in \mathbb{C}^n$ be a global optimizer of* (F2)*. Then the following holds:*

1. *$\mathcal{M}_{\mathcal{A}}(\mathbf{w}) = \mathcal{M}_{\mathcal{A}}(\mathbf{z})$ for all local minima $\mathbf{w}$ of* (F2)*.*

2. *The function f in* (F2) *has the strict saddle property.*

Theorem states that, with the conditions of α, β-stability and κ-cross-stability, the minimization problem (F2) has no spurious local minima, and any saddle point of the function f is *strict* in the sense that f has a strictly negative curvature at such a point. The latter property, called the strict saddle property, is defined in Section 3.6, where we also provide a formal statement of Theorem 3.6.7. Finally, based on the properties established about the loss landscape, we establish that a solution to problem (F2) can be obtained by applying a gradient based algorithm (from an arbitrary initial point), since such an algorithm is unlikely to converge to a saddle point.

Generally, designing the set of measurement matrices $\mathcal{A}$ to satisfy the properties can be quite non-trivial and may require area expertise. As a solution to this crisis, we focus our attention on random hermitian matrices and prove that all the necessary and sufficient conditions required for successful recovery of the unknown vector $\mathbf{z}$ are satisfied by random matrices in expectation. Further, we show the exact measurement complexity required for sub-gaussian random matrices to show such behavior with very high probability.

Our main result here states, states that the mapping $\mathcal{M}_{\mathcal{A}}$ is a near-isometry when the matrices A_i are chosen from a complex Gaussian distribution. A sketch of the statement is provided below.

Theorem 3.4.3 (sketch)**.** *Let $\mathcal{A} = \{A_i\}_{i=1}^m$ be a set of complex Gaussian Hermitian random matrices. Suppose that the number m of measurements satisfies $m > Cn$, for a large enough $C > 0$. Then, with a high probability, the following relation holds:*

1. *$\mathcal{M}_{\mathcal{A}}(\mathbf{w}) = \mathcal{M}_{\mathcal{A}}(\mathbf{z})$ for all local minima $\mathbf{w}$ of* (F2)*.*

2. *The function f in* (F2) *has the strict saddle property.*

In other words, we show that, with a sufficient number of random matrices $\{A_d\}$ and with high probability, the mapping $\mathcal{M}_{\mathcal{A}}$ nearly preserves distances with respect

to the distance measure defined in (3.2) as well as contains enough energy in entire $\mathbb{C}^n$ to be able to establish strict saddle property.

The formal statement along with the full proof is presented in Theorem B.4.3 in Section 3.5.

Such arguments have been used to prove crucial robustness results for the phase retrieval problem (however, only in $\mathbb{R}^n$); see e.g., Bandeira *et al.* (2014); Balan and Wang (2015); Candes *et al.* (2013). Our result for the optimization landscape of (F2) is in fact amenable to gradient based methods!

We would like to point out some nuances of what we have proved,

- (α, β)-stability, with $0 < \alpha \leq \beta$, is sufficient to guarantee the identifiability of solution $\mathbf{z}$ of (P1). This alone is however not sufficient to guarantee the recovery of $\mathbf{z}$ using gradient-based local methods by ℓ_2 loss minimization due to the possible existence of suboptimal local minimas.

- (α, β)-stability and κ-cross-stability, with $3\alpha > 2\beta$ and $\kappa \geq 0$, is sufficient for ensuring successful recover of $\mathbf{z}$ (F2) using local gradient-based methods.

- Sampling $\mathcal{A}$ from Guassian random hermitian matrices for $|\mathcal{A}| > Cn$ for sufficiently large $C > 0$ is sufficient to guarantee $3\alpha > 2\beta$ and $\kappa \geq 0$. Thereby being sufficient for solving (P1) through (F2)

(**Rewrite this para**) Most of the results related to Non-convex landscape and the strict saddle property of the optimization landscape provide guarantees only in high probability when the measuring vectors or matrices are sampled using random distributions. We are the first to isolate the properties of the measuring matrices $\mathcal{A}$ which are sufficient for obtaining strict saddle properties and identifiable proprieties. We also show Gaussian random matrices with enough measurement to satisfy these properties and hence satisfy the problem.

3.5 Identifiability of Quadratic Feasibility

The quadratic feasibility problem (P1) deals with recovering a complex vector $\mathbf{x} \in \mathbb{C}^n$ from the set of its quadratic measurements $\mathcal{M}_{\mathcal{A}}(\mathbf{x})$. To do so, we need to analyze the nature of the quadratic mapping $\mathcal{M}_{\mathcal{A}}(\cdot)$ and the role of the matrix ensemble $\mathcal{A}$ in influencing the same.

Note that given any complex matrix $A \in \mathbb{C}^{n\times n}$, note that $\langle A\mathbf{x}, \mathbf{x}\rangle = \langle A(-\mathbf{x}), (-\mathbf{x})\rangle$. In more generality, this phenomenon can be attributed to the symmetric nature of quadratic functions (or in general to even-powered homogeneous polynomials). The quadratic problem (P1) has the property of being unaltered to phase transitions of the unknown solution $\mathbf{x}$, i.e. if $\mathbf{x}$ is a solution to (P1) then $e^{i\theta}\mathbf{x},\ \theta \in [0, 2\pi]$ is a solution to (P1) as well. Hence the system of quadratic measurement in (P1) is fundamentally incapable of distinguishing up-to-a-sign ambiguity (in the real domain) and a phase factor (in the complex domain). To this end, we define $\mathbb{C}_{\angle} := \mathbb{C}\backslash \sim$ to be the phase ambiguous quotient space, where $\mathbf{x} \sim \mathbf{y}$ if and only if $\mathbf{x} = c\mathbf{y}$ for some $c \in \mathbb{C}$ with $|c| = 1$. We utilize $d_{\angle}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\mathbf{x}^\dagger - \mathbf{y}\mathbf{y}^\dagger\|_F$ for any $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ to be the phase invariant distance metric for this work. Notice that this $d_{\angle}(\cdot, \cdot)$ is invariant under phase shifts, i.e. $d_{\angle}(\mathbf{x}, \mathbf{y}) = 0$ if $\mathbf{x} = \mathbf{y}e^{i\theta}, \theta \in [0, 2\pi]$ and hence not a distance metric in $\mathbb{C}^n$. But $d_{\angle}(\cdot, \cdot)$ is a distance metric on $\mathbb{C}_{\angle}$. This concept of phase invariant distance metric is not novel and has been utilized in a lot of previous works Eldar and Mendelson (2014).

In light of the inability to distinguish between the phases of $\mathbf{x} \in \mathbb{C}^n$, it is impertinent to ask whether the system of quadratic measurements $\mathcal{M}_{\mathcal{A}}(\mathbf{x})$ is capable of recovering any unknown signal $\mathbf{x} \in \mathbb{C}^n$. Making this query mathematically precise, for all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, if $\mathbf{x} \nsim \mathbf{y}$ then can we be confident that $\mathcal{M}_{\mathcal{A}}(\mathbf{x}) \neq \mathcal{M}_{\mathcal{A}}(\mathbf{y})$? The answer to this is **Yes**. We refer to this property as *injectivity* of the mapping $\mathcal{M}_{\mathcal{A}}$ (for ease of

reference we sometimes refer to this as *injectivity* of matrix ensemble $\mathcal{A}$). To quantify this notion of *injectivity* of a given mapping $\mathcal{M}_{\mathcal{A}} : \mathbb{C}^n_{\angle} \to \mathbb{C}^m$ we define (α, β)-stability of mapping $\mathcal{M}_{\mathcal{A}}(\cdot)$ as follows,

Definition 3.5.1 ((α, β)-stability)**.** *Consider the mapping $\mathcal{M}_{\mathcal{A}}$ characterized by ensemble $\mathcal{A}$ as in* (3.1)*. We say that $\mathcal{M}_{\mathcal{A}}$ is (α, β)-stable w.r.t. metric $d_{\angle}(\cdot, \cdot)$, with $0 < \alpha \leq \beta$, if the following relation holds for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{C}^n$:*

$$\alpha d(\mathbf{x}_1, \mathbf{x}_2) \leq \|\mathcal{M}_{\mathcal{A}}(\mathbf{x}_1) - \mathcal{M}_{\mathcal{A}}(\mathbf{x}_2)\|_2 \leq \beta d(\mathbf{x}_1, \mathbf{x}_2). \tag{3.4}$$

The constants α, β depend on the choice of the matrix ensemble $\mathcal{A}$. The ratio of the constants α, β can also be thought of as a *condition number* or *distortion factor*, thereby allowing one to quantify the quality of the map; the higher the ratio between α and β, the better the ability of the mapping to distinguish between two distinct inputs. Our definition of (α, β)-stability requires both upper and lower bounds, which is instrumental in future results as we will see in the next section. This is in contrast to the stability concept considered for phase retrieval in Eldar and Mendelson (2014); Duchi and Ruan (2017). Similar notions of stability have been used to prove crucial robustness results for the phase retrieval problem (however, only in $\mathbb{R}^n$); see e.g., Bandeira *et al.* (2014); Balan and Wang (2015); Candes *et al.* (2013).

With this definition of stability in place, the next lemma establishes the connection between injectivity and stability of $\mathcal{M}_{\mathcal{A}}$.

Lemma 3.5.2. *The mapping $\mathcal{M}_{\mathcal{A}}$ is injective iff it is (α, β)-stable for some constants $0 < \alpha \leq \beta$.*

Complete proof of Lemma 3.5.2 can be found in Appendix B.1. As we will see in what follows, Lemma 3.5.2 allows one to assess the conditions under which the measurement model implied by the mapping $\mathcal{M}_{\mathcal{A}}$ is *identifiable*. Next, we start by

establishing necessary conditions on the quadratic mapping $\mathcal{M}_{\mathcal{A}}$ defined in (3.1) to ensure injectivity. Notice that the injectivity of the mapping is equivalent to the problem being identifiable (and hence solvable). The notion of injectivity and (α, β)-stability widely hinges on the nature of the ensemble $\mathcal{A}$. So how does one go about constructing such ensembles? Answer: Random ensemble.

In the following result, we prove that an ensemble of random matrices with sub-Gaussian entries suffices as $\mathcal{A}$. More precisely, a matrix $A \in \mathbb{C}^{n\times n}$ is a complex Hermitian sub-Gaussian matrix, if,

1. $\forall\ i,\ a_{ii} \sim \mathcal{SG}(0, \sigma^2)$.

2. $\forall\ i, j,\ i \neq j,\ a_{ij} \sim \mathcal{SG}(0, \frac{\sigma^2}{2}) + i\ \mathcal{SG}(0, \frac{\sigma^2}{2})$.

where $\mathcal{SG}(0, \sigma^2)$ indicates the sub-gaussian distribution with mean 0 and variance σ^2 and $\sim$ indicates "sampled from". An ensemble of complex Hermitian Gaussian random matrices $\mathcal{A} = \{A_d\}_{d=1}^m$ satisfies (α, β)-stability as follows,

Theorem 3.5.3. *Let $\mathcal{A} = \{A_i\}_{i=1}^m$ be an ensemble of complex sub-gaussian random matrices, and assume the number of measurements satisfies $m > Cn$ for some $C > 0$. Then, for any given $\xi \in (0, 1)$, there exist constants $C, \delta, \epsilon > 0$ such that, with probability at least $1 - \xi$, the following relation holds*

$$\alpha d(\mathbf{x}, \mathbf{y}) \leq \|\mathcal{M}_{\mathcal{A}}(\mathbf{x}) - \mathcal{M}_{\mathcal{A}}(\mathbf{y})\|_2 \leq \beta d(\mathbf{x}, \mathbf{y}).$$

where α, β are given by

$$\alpha \triangleq \frac{((1-2\delta)^2(1-\epsilon)}{(1+2\delta)^2}, \quad \beta \triangleq \frac{((1+2\delta)^2(1+\epsilon)}{(1-2\delta)^2}.$$

We refer the reader to Appendix B.4.2 for a proof of this lemma. The high probability argument is not new and has found application in several results; see e.g., Huang *et al.* (2019); Vershynin (2010); Meckes (2004).

Deterministic design of $\mathcal{A}$ so as to have (α, β)-stability can be a hard task and in some cases might require area expertise. While designing the measurement ensemble $\mathcal{A}$ is non-trivial, Wang and Xu (2017); Dan (2022) provide guidelines of the requirements for $\mathcal{A}$ to be injective. The results in Wang and Xu (2017) prove that a large class of matrix ensembles $\mathcal{A}$ does satisfy this injective property. In this next theorem, we focus on establishing sufficiency criteria for the mapping $\mathcal{M}_{\mathcal{A}}$ to be injective.

Theorem 3.5.4. *Consider $\mathcal{A} = \{A_i\}_{i=1}^m$ and the mapping $\mathcal{M}_{\mathcal{A}}$ be defined as in* (3.1). *Viewing $\{A_i\mathbf{x}\}_{i=1}^m$ as vectors in $\mathbb{R}^{2n}$, we can say the following:*

1. *$\mathcal{M}_{\mathcal{A}}$ is injective.*
2. *$\forall I \subset [m]$, then either $dim(span(\{A_i\mathbf{x}\}_{i\in I}))\geq 2n-1$ or $dim(span(\{A_i\mathbf{x}\}_{i\in I^c}))\geq 2n-1$ for all vectors $\mathbf{x} \in \mathbb{C}^n\backslash 0$ (complementary property Balan et al. (2006))*
3. *$\forall I \subset [m]$, then either $span(\{A_i\mathbf{x}\}_{i\in I}) = span(\{i\mathbf{x}\})^{\perp}$ or $span(\{A_i\mathbf{x}\}_{i\in I^c}) = span(\{i\mathbf{x}\})^{\perp}$ for all vectors $\mathbf{x} \in \mathbb{C}^n\backslash 0$.*

Given an ensemble $\mathcal{A}$, Theorem 3.5.4 states that as long as the mapping $\mathcal{M}_{\mathcal{A}}$ satisfies the complementary property Balan *et al.* (2006), we can be assured that $\mathcal{M}_{\mathcal{A}}$ is injective. Please refer to Appendix B.1 for the complete proof. Even though $\alpha > 0$ is a necessary and sufficient condition for (P1) to be solvable, when dealing with data-analytic methods in the presence of noise, the more the disparity between α and β values, the harder it gets to provide guarantees on obtaining the unknown signal $\mathbf{x}$.

3.5.1 Problem Complexity : $\mathbb{R}^n$ vs $\mathbb{C}^n$

The quadratic feasibility problem (P1) has vastly different behavior when the problem is set up in $\mathbb{R}^n$ or $\mathbb{C}^n$. Firstly consider the well-studied problem of phase retrieval, a special case of quadratic feasibility when the ensemble $\mathcal{A}$ is a collection of

rank-1 matrices, i.e. $\mathcal{A} = \{\mathbf{a}_i\mathbf{a}_i^\dagger\}_{i=1}^m$ (when considering $\mathbb{C}^n$) or $\mathcal{A} = \{\mathbf{a}_i\mathbf{a}_i^\top\}_{i=1}^m$ (when considering $\mathbb{R}^n$). The authors in Botelho-Andrade *et al.* (2016) argue the fundamental difference between the two concepts: *phase retrieval* and *phaseless recovery* (please refer Botelho-Andrade *et al.* (2016) for more details). They prove that these two concepts coincide when considering the phase retrieval in $\mathbb{R}^n$ but are different in $\mathbb{C}^n$. The solution space of Problem (P1) is isolated solutions $\mathbf{z}$ or $-\mathbf{z}$ in the case of $\mathbb{R}^n$, but a continuum $[e^{i\theta}\mathbf{z}, \theta = [0, 2\pi)]$ in the case of $\mathbb{C}^n$. A fundamental lower bound on the cardinality of $\mathcal{A}$ is established to be $2n-1$ for $\mathbb{R}^n$ but is still an open question for $\mathbb{C}^n$ (some results can be seen in Heinosaari *et al.* (2013); Botelho-Andrade *et al.* (2016)).

Authors in Wang and Xu (2017) prove that, when the setup (P1) is considered in $\mathbb{R}^n$, the set $\{\mathbf{x} : \mathcal{M}(\mathbf{x}) = 0\}$ is discrete. We dedicate Appendix B.2 to developing Lemma B.2.2 which mathematically proves that this is not the case for (P1) in $\mathbb{C}^n$ and the set of $\{\mathbf{x} : \mathcal{M}(\mathbf{x}) = 0\}$ is a continuum, further depicting the dissimilarity and the need to analyze $\mathbb{R}^n$ and $\mathbb{C}^n$ separately.

3.6 Solving Quadratic Feasibility

Having established conditions under which the mapping $\mathcal{M}_\mathcal{A}$ represents an identifiable measurement model and hence (P1) is solvable, we next focus on tractable approaches for the retrieval of the unknown signal $\mathbf{z} \in \mathbb{C}^n$. Consider the ℓ_2-loss formulation of problem (P1),

$$\min_{\mathbf{x}\in\mathbb{C}^n} f(\mathbf{x}), \qquad f(\mathbf{x}) \triangleq \frac{1}{m}\sum_{i=1}^{m} |\langle A_i\mathbf{x}, \mathbf{x}\rangle - c_i|^2 . \tag{F2}$$

The reader can easily reason that the above loss function is non-convex (being 0 at the $\mathbf{x} = \pm\mathbf{z}$ but > 0 at $\mathbf{x} = 0$). This negates a lot of the existing literature on gradient descent and its convergence guarantees on convex functions. Naively attempting to solve (F2) with gradient descent can lead to convergence to local minima, thereby

failure in retrieving the unknown signal $\mathbf{x}$.

It is a well-known result that, for nonconvex loss functions, first-order gradient-based approaches need not converge to a global minimum. This is one of the central challenges stopping the widespread application of local methods for tackling non-convex problems. One of the primary attributes of this difficulty is the inability to distinguish local from global minima due to a lack of global information. Hence retrieving the generator $\mathbf{z}$ through the ℓ_2-loss minimization problem (F2) is non-trivial. In order to ensure that gradient-based algorithms can recover the unknown solutions we define a sufficiency condition for the matrix ensemble $\mathcal{A}$ to ensure retrievability for local gradient-based methods using (F2).

Lately, nonconvex optimization has received considerable attention due to the abundant practical applications and recent advances in tackling a subset of this problem category. In particular, methods like SGD and other gradient-based methods have been shown to be astonishingly successful in converging to global minima in many nonconvex problems Du *et al.* (2018); Chen *et al.* (2018); Jin *et al.* (2017). Arguably, the reason for this is that the optimization landscape of these (somewhat well-behaved) nonconvex problems enjoy advantageous properties which are crucial in the empirical success of the gradient based methods Bower *et al.* (2018); Davis *et al.* (2017); Sun *et al.* (2018). The work in Sun *et al.* (2018) proves that the ℓ_2-loss function for the phase retrieval problem enjoys properties such as all local minima being global, and each saddle point having a strictly negative curvature.

Typically such landscape properties are proved in population studies and the success of gradient descent is proved as a result of the concentration phenomenon. While random matrix ensembles are alluring choices for setting up (P1), in many applications, the setup in (P1) reflects the physical constraints in the physical world which may not be justified to be modeled by Gaussian random matrices. Hence in the

rest of the section, we focus on *characterizing* the relevant landscape properties of the ensemble $\mathcal{A}$ so as to enable all local gradient methods to reach global minima. Note that, under the right condition, our analysis obviates the need to provide initialization to first-order gradient-based algorithms Huang *et al.* (2019); Chen *et al.* (2023a).

3.6.1 Conditioning on $\mathcal{A}$

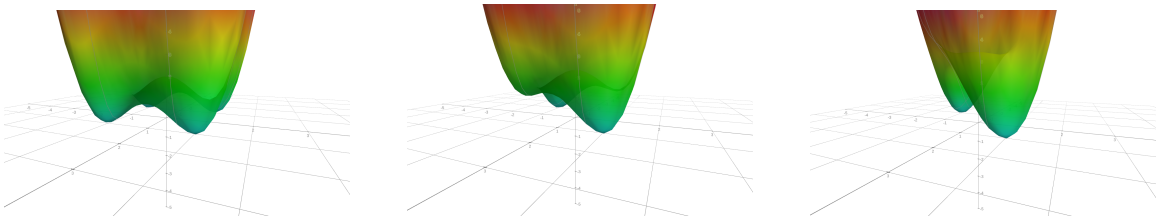
We provide an intuitive example in $\mathbb{R}^n$ to make the reader understand better. Consider the toy problem of retrieving the signal $\mathbf{z} = [-1, 1] \in \mathbb{R}^2$ as in (P1) with the following measuring matrix ensemble $\mathcal{A}_a$ (parameterized by $a \in \mathbb{R}^+, 0 < a < 1$),

$$\mathcal{A}_a = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 2a \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 2a & 0 \end{bmatrix} \right\} \tag{3.5}$$

Please note that $\mathcal{A}_a$ satisfies $(\min\{2a, 1\}, \max\{2a, 1\})$-stability. The ℓ_2-loss formulation can be given by,

$$f_a\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \frac{1}{4}\left[(x_1^2 - 1)^2 + (x_2^2 - 1)^2 + 2(2ax_1x_2 + 2a)^2 \right] \tag{3.6}$$

Given $a > 0$, it is easy to verify that $f_a(\mathbf{x}) = 0$ if and only if $\mathbf{x} = [1, -1]$ or $\mathbf{x} = [-1, 1]$ and $f_a(\mathbf{x}) > 0$ for all other $\mathbf{x} \in \mathbb{R}^2, a > 0$. The following graphs show the landscape of $f_a(\cdot)$ for different values of a:



(a) $a = 0.05$ (b) $a = 0.2$ (c) $a = 0.6$

Figure 3.1: Function landscape of f_a for different values of a

As can be observed, increasing the value of a helps in elevating sites of potential local minima (these being $[1,1]$ and $[-1,-1]$). To solidify this understanding, we analyze it from the Hessian point of view,

$$\nabla^2 f_a\left(\begin{bmatrix}1\\1\end{bmatrix}\right) = \nabla^2 f_a\left(\begin{bmatrix}-1\\-1\end{bmatrix}\right) = \begin{bmatrix}2+4a^2 & 12a^2\\ 12a^2 & 2+4a^2\end{bmatrix} \tag{3.7}$$

The eigenvalues can be computed to be $\lambda_1 = 2(8a^2+1), \lambda_2 = -2(4a^2-1)$. Note that $\lambda_2 > 0 \;\; \forall a < \frac{1}{2}$ and $\lambda_2 < 0 \;\; \forall a > \frac{1}{2}$. Thus, even if $\mathcal{A}_a$ is $(\min\{2a,1\}, \max\{2a,1\})$-stable, if $a < \frac{1}{2}$, $\nabla f(\mathbf{x}) \succ 0$ for $\mathbf{x}$ all $\mathbf{x} = \begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}-1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}, \begin{bmatrix}-1\\-1\end{bmatrix}$. Note that, for $a > \frac{1}{2}$, $\nabla f(\mathbf{x}) \not\succ 0$ (since $\lambda_1 > 0$ but $\lambda_2 < 0$) for $\mathbf{x} = \begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}-1\\-1\end{bmatrix}$ hence not local minima, but saddle points with strict negative curvature.

3.6.2 Phase Irregularity Correction

Following up on the intuition illustrated in Toy-example in Section 3.6.1, we concretize the criteria to make sure ℓ_2-loss function does not induce local minima on any $\mathbf{x}$ apart from the solution to Problem (P1). We first formalize the definition of the class of points which could potentially lead to situations as expressed in the toy example.

Definition 3.6.1 (Phase-irregular vectors)**.** *Vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ are phase-irregular if and only if $\nexists\; a \in \mathbb{C}$ such that $\mathbf{x} = a\mathbf{y}$.*

Intuitively, Definition 3.6.1 formalizes the notion of distinguishing two vectors in $\mathbf{C}^n$ *only* based on *irregular* phase changes between the individual dimensions of the vectors. Note that Definition 3.6.1 displays a notion complimentary to "Unique up-to phase ambiguity" when restricted to norm constrained $\mathbf{x}$. To give a more intuitive

example, let $\mathbf{x} = [x_1, x_2, \dots], \mathbf{y} = [y_1, y_2, \dots] \in \mathbb{C}^n$ be such that $y_i = x_i \ \ \forall i \in [2, n]$ and $y_1 = e^{i\theta} x_1$ for some $\theta \in [0, 2\pi]$. Then $\mathbf{x}$ and $\mathbf{y}$ are phase-irregular. Next we define a distance notion $d_\varphi(\cdot, \cdot)$ to capture the phase irregularity as follows,

$$d_\varphi(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\mathbf{y}^\top - \mathbf{y}\mathbf{x}^\top\|_F \tag{3.8}$$

Note the following properties of $d_\varphi(\cdot, \cdot)$,

1. $d_\varphi(\mathbf{x}, \mathbf{y}) > 0$ is and only if $\mathbf{x}, \mathbf{y}$ are phase-irregular.
2. $d_\varphi(\mathbf{x}, a\mathbf{y}) = |a| d_\varphi(\mathbf{x}, \mathbf{y})$ for any $a \in \mathbb{C}$.
3. $d_\varphi(\mathbf{x}, \mathbf{x} + \mathbf{y}) = d_\varphi(\mathbf{x}, \mathbf{y})$
4. $d_\varphi(\mathbf{x}, \mathbf{y} + \mathbf{z}) \leq d_\varphi(\mathbf{x}, \mathbf{y}) + d_\varphi(\mathbf{x}, \mathbf{z})$

The notion of phase-irregular distance $d_\varphi(\cdot, \cdot)$ ties well with Toy example in Section 3.6.1, as the $d_\varphi([1, -1], [1, 1]) > 0$ since $[1, -1], [1, 1]$ are phase-irregular. Next, we define the condition for our ensemble $\mathcal{A}$ to incorporate $d_\varphi(\cdot, \cdot)$ termed as κ-*phase discriminating*, in order to ensure elimination of the suboptimal local minima emerging due to $\mathbf{x}$ which are phase-irregular with solution of (P1),

Definition 3.6.2 (κ-phase discriminating)**.** *For any vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ and $\kappa > 0$, we say that the matrix ensemble $\mathcal{A}$ is said to be κ-phase discriminating if*

$$\begin{aligned} &\frac{1}{m} \sum_{d=1}^{m} \left(\langle A_d, \mathbf{y}\mathbf{x}^\dagger \rangle \langle A_d, \mathbf{x}\mathbf{y}^\dagger \rangle - \langle A_d, \mathbf{y}\mathbf{y}^\dagger \rangle \langle A_d, \mathbf{x}\mathbf{x}^\dagger \rangle \right) \\ &\geq \kappa d_\varphi(\mathbf{x}, \mathbf{y})^2. \end{aligned} \tag{3.9}$$

Generalizing from the toy setup in subsection 3.6.1, the potential sites of saddle point/local minima (with appropriate assumptions on a) were precisely the locations in $\mathbb{C}^n$ where only a few dimensions have a phase shift. Intuitively, definition (3.9)

ensures that the ensemble $\mathcal{A}$ eliminates these locations of potential local minima, by making sure they behave as a saddle point (here $\kappa > 0$ behaves as $a > \frac{1}{2}$ in our toy example). The concern at this point would be the practicality of κ - phase-discriminating (Definition 3.6.2). In the next result we show that such a condition is satisfied by large class of ensemble $\mathcal{A}$,

Theorem 3.6.3. *Let $\mathcal{A} = \{A_i\}_{i=1}^m$ be an ensemble of complex sub-gaussian random matrices, and assume the number of measurements satisfies $m > Cn$ for some $C > 0$. Then, for any given $\xi \in (0,1)$, there exist constants $C, \delta, \epsilon > 0$ such that, with probability at least $1 - \xi$, the following relation holds*

$$\begin{aligned}&\frac{1}{m}\sum_{d=1}^{m}\left(\langle A_d, \mathbf{y}\mathbf{x}^\dagger\rangle\langle A_d, \mathbf{x}\mathbf{y}^\dagger\rangle - \langle A_d, \mathbf{y}\mathbf{y}^\dagger\rangle\langle A_d, \mathbf{x}\mathbf{x}^\dagger\rangle\right)\\ &\geq \frac{((1-2\delta)^2(1-\epsilon)}{(1+2\delta)^2} d_\varphi(\mathbf{x},\mathbf{y})^2.\end{aligned}$$

Our criteria in this section was to focus on phase irregularity correction and hence out choice of utilizing $d_\varphi(\cdot,\cdot)$ over $d_\angle(\cdot,\cdot)$ (Note that for any scalar number a, $|a| \neq 1$, $d_\varphi(\mathbf{x}, a\mathbf{x}) = 0$ but $d_\angle(\mathbf{x}, a\mathbf{x}) \neq 0$ showcasing the focus of $d_\varphi(\cdot,\cdot)$ on only phase irregularity detection)

Using the notions of α, β-stability and κ-phase discriminating, we provide guarantees on the landscape of the ℓ_2-loss function (F2) in the next subsection.

3.6.3 Benign Landscape Guarantees

While the toy setup in Section 3.6.1 is very limited in terms of its construction, intuitively it suggests that under certain conditions on matrix ensembles $\mathcal{A}$, the landscape of Equation (F2) might be "nice" enough to enable gradient descent methods to recover the unknown signal $\mathbf{x}$. In this subsection, we go deeper and establish these conditions concretely for ℓ_2-loss function (F2)

Observe that the ℓ_2-loss function (F2) $f : \mathbb{C}^n \to \mathbb{R}$ is not differentiable in the complex space $\mathbb{C}^n$. Hence it is challenging to address the problem (F2) in a standard analysis technique of gradient descent. Instead, we utilize techniques from Wirtinger calculus Candes *et al.* (2015). Using this, our first step is to define the notion of a strict saddle function.

Definition 3.6.4. *A function f is said be strict saddle function if for any $\mathbf{x} \in \mathbb{C}^n$, at least one of the following statements is true:*

1. $\|\nabla f(\mathbf{x})\| > 0$;
2. $\begin{bmatrix} \mathbf{z} \\ \bar{\mathbf{z}} \end{bmatrix}^\dagger \nabla^2 f(\mathbf{x}) \begin{bmatrix} \mathbf{z} \\ \bar{\mathbf{z}} \end{bmatrix} < 0$ *for some* $\mathbf{z} \in \mathbb{C}^n$;
3. $\mathbf{x}$ *is the global minimum, i.e. if* $\mathbf{x} \in \mathbb{C}^n$ *satisfies* $\nabla f(\mathbf{x}) = 0$ *and* $\nabla^2 f(\mathbf{x}) \succeq 0$ *then* $\mathcal{M}_{\mathcal{A}}(\mathbf{x}) = \mathcal{M}_{\mathcal{A}}(\mathbf{y})$ *where* $\mathbf{y}$ *is the unknown signal in* (P1).

where the notion of gradient $\nabla f(\cdot)$ and hessian $\nabla^2 f(\cdot)$ are derived through wirtinger calculus.

Intuitively, this implies that every $\mathbf{x} \in \mathbb{C}^n$ either violates optimality (condition 1 and 2) or is a global optimum. A line of recent work Lee *et al.* (2016, 2017); Jain *et al.* (2017) has explored the efficacy of gradient based methods in finding a local optimum of functions satisfying Definition 3.6.4.

We analyze the optimization landscape of (P1). For a mild condition on the stability parameters α, β and κ we show that *every local minimum is in fact global* (upto the equivalence relation $\sim$). Our next main result states that the function f in (F2) is strict saddle.

Theorem 3.6.5. *Let the scalar vector $\mathbf{c} \in \mathbb{R}^n$ be generated by quadratic measurements of an unknown vector $\mathbf{z} \in \mathbb{C}^n$ characterizing the measurements used in the objective*

function f of problem (F2). *Let $\mathcal{A}$ be a set of measurement matrices which satisfy (α, β)-stability with $2\beta < 3\alpha$ and κ-phase-disciminating with $\kappa \geq 0$. Then, the following statements hold:*

1) The function f is strict saddle, and

2) Every local minimum $\mathbf{w}$ of f satisfies $d_{\angle}(\mathbf{w}, \mathbf{z}) = 0$, where $\mathbf{z}$ is the global minima.

Proof sketch. Notice that to show that the function f in (P2) is a strict saddle function, it suffices to only consider the points $\mathbf{w} \in \mathbb{C}^n$ such that $\|\nabla f(\mathbf{w})\| = 0$ (otherwise, 1) of Definition 3.6.4 is satisfied). For all such points, we analyze the behavior of the Hessian and establish that there exists a direction $\Delta \in \mathbb{C}^n$ such that the following inequality holds :

$$\begin{bmatrix} \Delta \\ \bar{\Delta} \end{bmatrix}^{\dagger} \nabla^2 f(\mathbf{w}) \begin{bmatrix} \Delta \\ \bar{\Delta} \end{bmatrix} < 0, \tag{3.10}$$

whenever $d_{\angle}(\mathbf{x}, \mathbf{w}) \neq 0$ and $\nabla f(\mathbf{w}) = 0$.

This implies that there is a direction where the hessian has a strict negative curvature, and hence $\mathbf{w}$ cannot be a minima. In other words, we can conclude that: (1) All local minima satisfy $d(\mathbf{x}, \mathbf{z}) = 0$, and (2) all saddle points have a strictly negative curvature. □

Theorem 3.6.5 is the first result we know of which extracts conditions on measuring matrices $\mathcal{A}$ such that the loss functions have benign landscape properties. Most of the results in terms of proving good landscape properties related to nonconvex functions deal with high probability results in the presence of randomness. We establish conditions that can be used as a design protocol for deterministic measuring matrices $\mathcal{A}$.

While none of the works have concretely stated the precise condition as we have in Theorem 3.6.5, a lot of works in literature have used $3\alpha > 2\beta$ internally to prove results Ge *et al.* (2017); Sun *et al.* (2018). Note that $\kappa \geq 0$ is a requirement only when analyzing Problem (P1) in the complex space $\mathbb{C}^n$. The Authors in Ge *et al.* (2017) could prove results for similar problems, without the need to invoke such property.

Finally, we remark that the properties of the optimization landscape that we have established allow one to use any gradient based iterative method to find a global optimum of the problem (P2) – hence, find a solution to the quadratic feasibility problem. Furthermore, our results above also imply that a gradient method, with an arbitrary initial point, would work, which is in sharp contrast with the existing works, such as Huang *et al.* (2019). Formally, we have the following result.

Corollary 3.6.6. *Consider a gradient method applied to minimize the function f in* (F2)*. Then, for an arbitrary initial point, the method point converges to a global minimum of the loss function f associated with the quadratic feasibility problem.*

Given the landscape properties we have derived in Theorem 3.6.5, this result follows in a straightforward manner, for instance, from Theorem 4.1 in Lee *et al.* (2016). We would like to remark here that the broad flow of ideas in our proof of Theorem 3.6.7 bears similarities to those in papers like Ge *et al.* (2017); Bhojanapalli *et al.* (2016b). However, to the best of our knowledge, the present paper is the first to derive such results based on the characterization of ensemble $\mathcal{A}$ in the complex domain. One of the novel contributions of the paper is the isolation of the two conditions namely, one which provides identifiablity to quadratic feasibility problems and another which lets it be solved by computationally efficient algorithms. An interesting observation is that in most of the previous works Eldar and Mendelson (2014); Sun *et al.* (2018) both of these conditions are satisfied either by construction or by randomness and hence were

never highlighted. In this work, we also focus on the necessary conditions required for designing measurement matrices so as to solve (F2) by local-gradient-based methods.

We analyze the optimization landscape of (P2) when our measurement matrices are Hermitian and complex Gaussian, and show that with a high probability *every local minimum is in fact global* (upto the equivalence relation $\sim$). Our next main result states that the function f in (P2) is the strict saddle.

Theorem 3.6.7. *Let $\{A_i\}_{i=1}^m$ be a set of complex $n \times n$ Gaussian random matrices, and let $m > Cn$ for some constant $C > 0$. Let the scalars $\{c_i\}_{i=1}^m$ characterizing the objective function f of problem (P2) be generated by quadratic measurements of an unknown vector $\mathbf{z}$. Then, for any given $\xi \in (0, 1)$, there exist positive constants β, γ, and ζ such that the following statements hold with probability at least $1 - \xi$:*

1) The function f is (β, ζ, γ)-strict saddle, and

2) Every local minimum $\mathbf{w}$ of f satisfies $d(\mathbf{w}, \mathbf{z}) = 0$

where (β, ζ, γ)-strict saddle is an extension of Definition 3.6.4 (more on this in Appendix B.5). Note a crucial difference between Theorem 3.6.5 and Theorem 3.6.7 : While Theorem 3.6.7 holds with probability $1 - \xi$, Theorem 3.6.5 does not need such a condition. This is because in Theorem 3.6.5, $3\alpha > 2\beta$ and $\kappa \geq 0$ holds with certainty and the same holds with high probability in Theorem 3.6.7. This showcases the fundamental importance of (α, β)-stability and κ-phase discriminating property in proving landscape results.

3.7 Robustness Analysis

The results in Theorem 3.6.5 & Theorem 3.6.7 hold only when the measurement process is noiseless. This rarely holds in the real-world measurement processes and

hence this section is devoted to the analysis of retrieving the unknown $\mathbf{x}$ from noisy measurements. Consider the following corrupted problem definition:

$$\begin{aligned} &\text{find} \quad \mathbf{x} \\ &\text{such that } \langle A_i\mathbf{x}, \mathbf{x}\rangle = c_i + \eta_i, \quad \forall i = 1, 2, \ldots, m \end{aligned} \tag{P_η}$$

where the measurements c_i are generated from an unknown $\mathbf{z} \in \mathbb{C}^n$, i.e. $c_i = \langle A_i\mathbf{z}, \mathbf{z}\rangle$ and $\eta_i \sim \mathcal{N}(0, \sigma^2) \quad \forall i \in [n]$ are i.i.d. Gaussian noise variables.

Due to the presence of Gaussian noise, an ideal method to find an estimate $\hat{\mathbf{x}}$ is by minimizing the following ℓ_2-loss function,

$$\min_{\mathbf{x}\in\mathbb{C}^n} g_\eta(\mathbf{x}) \triangleq \frac{1}{m}\sum_{i=1}^{m} \left|\langle A_i\mathbf{x}, \mathbf{x}\rangle - c_i - \eta_i\right|^2. \tag{F_η}$$

To this end, our next result bounds the distance between the generator $\mathbf{z}$ of the measurements c_i and the solution of the problem (F_η).

Theorem 3.7.1. *Let $\{A_i\}_{i=1}^m$ be a set of complex $n \times n$ Gaussian random matrices, and let $m > Cn$ for some large constant $C > 0$. Let the scalars $\{c_i\}_{i=1}^m$ characterizing the objective function f of problem P1 be generated by quadratic measurements of an unknown vector $\mathbf{z}$. Let the Let $\hat{\mathbf{x}} \in \mathbb{C}^n$ be such that $\|\nabla g_\eta(\hat{\mathbf{x}})\| \leq \delta$, where g_η is the ℓ_2-loss function (F_η). Then with probability $1 - c_1e^{-c_2m\epsilon} - 2e^{-c_3m\epsilon_\eta}$ the following holds:*

$$\|\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^*\|_F \leq \frac{4(\delta + \epsilon_\eta\sigma_\eta)}{1 - 5\epsilon} \tag{3.11}$$

where $\sigma_\eta \in \mathbb{R}$ is the noise variance and $c_0, c_1, c_3 > 0$ are some constants.

The above theorem proves that in the presence of Gaussian noise, the actual solution of the problem (P1) is in close vicinity of the solution obtained by solving the noisy ℓ_2-loss function (F_η). For the complete proof of the above Theorem, please refer to the appendix

Chapter 4

BANDIT-BASED MULTI-AGENT SEARCH

4.1 Abstract

Autonomous search using teams of multiple agents needs tractable coordination strategies between the search agents. The strategy must lower the time to identify interesting areas in the search environment, lower the costs/energy usage by the search agents during movement and sensing, and be resilient to the noise present in the sensed data due to the use of low-cost and low-weight sensors. We propose a data-driven, multi-agent search algorithm to achieve these goals using the framework of thresholding multi-armed bandits. For our algorithm, we also provide finite upper bounds on the time taken to complete the search, on the time taken to label all interesting cells, and on the economic costs incurred during the search.

4.2 Introduction

Autonomous multi-agent search for objects/phenomena of interest over large areas is crucial in several applications, including environmental monitoring, agriculture, search-and-rescue, and wildlife monitoring. Given a grid environment to search, we study the problem of identifying *all* interesting cells (cells that contain an object/phenomenon of interest) using multiple search agents, each equipped with a noisy sensor. We require the search agents to satisfy multiple requirements. First, the search agents must coordinate and quickly identify interesting cells, which is essential in time-sensitive applications like search-and-rescue. Additionally, they must minimize economic costs associated with the search, which could include the energy used by the search agents

due to movement and sensing. Finally, they must make decisions on locations to sense based on noisy observations obtained online from low-cost and low-weight sensors typical of such search systems. In this paper, we propose *a data-driven, multi-agent search algorithm that addresses these requirements using the framework of thresholding bandits.*

A solution to the search problem is the *label-then-move search* (see (Rolf *et al.*, 2021) for a variant of this search). In this search strategy, we partition the search space into disjoint sets of grid cells that are assigned to the search agents. Each search agent starts at some cell within its assigned set of cells, collects enough data at a grid cell until it is confident enough to label the grid cell as interesting or uninteresting, and then moves on to another grid cell within its assigned set. The label-then-move search strategy ignores the data collected online to decide on the next location to sense. Consequently, it can spend a significant amount of time in labeling uninteresting cells and may not be well suited for time-sensitive applications.

Recently, multi-armed bandits (MAB) have also been proposed for the multi-agent search problem. Recall that MAB is a special class of reinforcement learning algorithms where the current actions *do not* impact future reward (Lattimore and Szepesvári, 2020). MAB-based algorithms typically enjoy non-asymptotic guarantees of performance with minimal assumptions, unlike general reinforcement learning algorithms (Lattimore and Szepesvári, 2020). (Rolf *et al.*, 2021; Du *et al.*, 2021) propose MAB-based search strategies that identify the maximal or top-k interesting cells in a grid and require prior knowledge of the number of interesting cells. Instead, our work focuses on identifying *all* interesting grid cells without prior knowledge of the total number of interesting cells.

In this paper, we combine the label-then-move search and thresholding MAB-based search strategies, as illustrated in Figure 4.1. Similarly to (Locatelli *et al.*,

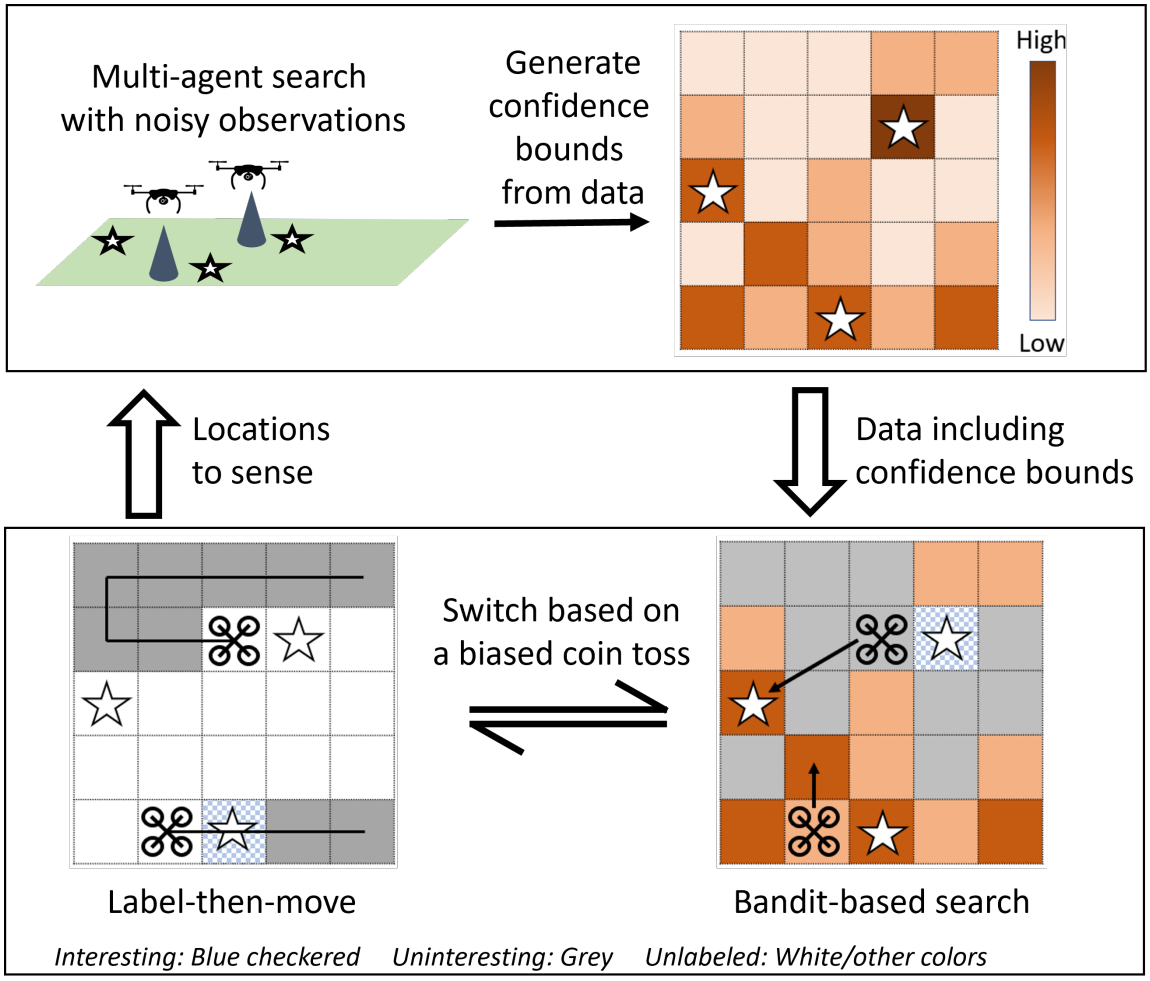


Figure 4.1: Data-driven multi-agent search under noisy observations. The proposed approach switches between a bandit-based search and a *label-then-move search* with a user-specified probability. The bandit-based search optimizes a surrogate function constructed using noisy observations for making decisions on locations to sense. The *label-then-move search* makes the agents follow a fixed, pre-determined pattern independent of the data collected online.

2016; Mason *et al.*, 2020), we decide on the next locations to sense by maximizing a surrogate function that is constructed using the data collected online. While a vanilla application of the thresholding bandit-based approach is near-optimal in terms of search completion time (Locatelli *et al.*, 2016), it may have high economic costs. A key feature of our approach is the ability to trade off the desire for efficient search with the need to lower the economic costs associated with the search.

Several other strategies have also been proposed for multi-agent search (Drew, 2021; Queralta *et al.*, 2020). Popular approaches include algorithms based on submodular maximization (Krause and Guestrin, 2007), algorithms combining Voronoi-based search (Bullo *et al.*, 2009) with function approximation (Schwager *et al.*, 2015; Luo *et al.*, 2019), active sensing/perception algorithms (Bajcsy *et al.*, 2018), graph-based search algorithms (Kapoutsis *et al.*, 2017; Best *et al.*, 2018), and algorithms based on statistical learning (Marchant and Ramos, 2012; Ghods *et al.*, 2021; Banerjee *et al.*, 2022). However, these works may require perfect sensing, may not have finite-time guarantees on the search performance, and/or may have high economic costs of search associated with movement and sensing.

The main contributions of this paper are: 1) to propose a multi-agent search algorithm that accommodates noisy observation data using a combination of thresholding MAB and label-then-move search, and 2) to characterize the performance of the algorithm by determining finite upper bounds on the time taken to complete the search, time taken to label all interesting cells, and the economic costs incurred during the search by our algorithm. We propose two metrics, *priority labeling time* and *economic cost*, to study the performance of the proposed algorithm. Additionally, with respect to the existing literature in MAB, we integrate coordination requirements and the physical limitations of switching actions directly into the algorithm. Finally, we demonstrate the efficacy of our approach in numerical simulations.

4.3 Problem Formulation

Let $\mathcal{G}$ denote the set of grid cells defining the search environment. Let the autonomous search team have d agents, each equipped with a noisy sensor. In the autonomous multi-agent search problem, we must identify the sequence of grid cells to visit and sense based on the noisy sensor data collected online, and return a set of interesting grid cells. We refer to a cell as *interesting* if it contains an object/phenomenon of interest. In this section, we formalize the autonomous multi-agent search problem as a variant of the thresholding MAB, characterize suitable metrics to analyze the performance of a search algorithm and state the problems addressed in this paper.

MAB formulation of the multi-agent search problem:

We cast the search problem as a $|\mathcal{G}|$-armed bandit problem where the arms are grid cells with $\mathcal{G} \triangleq \{1, 2, 3 \ldots, |\mathcal{G}|\}$, and $|\mathcal{G}|$ denotes the cardinality of the set $\mathcal{G}$. Let $\mathcal{S}^{\complement} = \mathcal{G} \setminus \mathcal{S}$ be the complement of any set $\mathcal{S} \subseteq \mathcal{G}$.

At each time step, the search team of d agents selects a set of d *distinct* grid cells to visit simultaneously. Each visit to a grid cell returns a binary indication of whether the cell is interesting. However, the observation data may be corrupted by noise, arising from sensing limitations and perception errors. Formally, a visit to grid cell $i \in \mathcal{G}$ results in a draw of a sample from a corresponding Bernoulli random variable ν_i with mean μ_i. The mean μ_i is influenced by the underlying spatial distribution of the interesting cells, the characteristics of the noisy sensors, and the perception algorithms used by the agents. We assume that the Bernoulli random variables for any two cells in $\mathcal{G}$ are mutually independent.

Remark 4.3.1. *We do not assume prior knowledge of μ_i for any grid cell $i \in \mathcal{G}$ or*

the total number of interesting cells.

Desired outcome of the search:

For a user-specified *threshold* $\theta \in (0, 1)$, upon the completion of the search, we seek to identify the set of grid cells,

$$\mathcal{S}_\theta = \{i \in \mathcal{G} | \mu_i \geq \theta\} \subset \mathcal{G}. \tag{4.1}$$

The set $\mathcal{S}_\theta$ is the set of grid cells that may be sensed as interesting with a probability of at least θ.

We make the following assumption to obtain finite-sample guarantees despite the noisy sensors on the search agents.

Assumption 4.3.2 (Labeling error tolerance)**.**
The labeling error for grid cells $i \in \mathcal{G}$ with $\mu_i \in (\theta - \epsilon, \theta + \epsilon)$ may be ignored for some (small) tolerance $\epsilon > 0$.

Assumption 4.3.2 is motivated by the observation that deciding if $|\mu_i - \theta| > \epsilon$ for any grid cell $i \in \mathcal{G}$ using a finite number of samples becomes harder as μ_i approaches θ (Jun *et al.*, 2016; Lattimore and Szepesvári, 2020). Under Assumption 4.3.2, any set $\mathcal{K} \subseteq \mathcal{G}$ that satisfies

$$\mathcal{S}_{\theta+\epsilon} \subseteq \mathcal{K} \subseteq \mathcal{S}_{\theta-\epsilon}, \tag{4.2}$$

is an acceptable approximation of $\mathcal{S}_\theta$.

We use the notion of a *search policy* to characterize a multi-agent search.

Definition 4.3.3 (Search policy)**.** *Let $\mathcal{H}(\tau) = \{\mathcal{H}_i(\tau)\}_{i \in \mathcal{G}}$, where $\mathcal{H}_i(\tau)$ is the history of observations at grid cell $i \in \mathcal{G}$ collected by all agents until time τ. Let $\pi_\tau : \mathcal{H}(\tau) \to \mathcal{G}^d$ be a function that maps $\mathcal{H}(\tau)$ to d distinct grid cells in $\mathcal{G}$. We define a*

search policy $\boldsymbol{\pi}_t$ as a sequence of functions $\boldsymbol{\pi}_t = \{\pi_\tau\}_{0\leq\tau\leq t}$. We will drop the subscript on $\boldsymbol{\pi}_t$ when time is not relevant.

Performance metrics: Let the multi-agent search using the search policy $\boldsymbol{\pi}$ terminates at time step $T_{\boldsymbol{\pi}} \in \mathbb{N}$, and return a sequence of sets $\{\mathcal{K}(t)\}_{t=0}^{T_{\boldsymbol{\pi}}}$ with $\mathcal{K}(t) \subseteq \mathcal{G},\ \forall t$. A successful search policy $\boldsymbol{\pi}$ has a low *labeling error* upon termination, i.e., it satisfies

$$\mathbb{P}\left[(\mathcal{S}_{\theta+\epsilon} \setminus \mathcal{K}(T_{\boldsymbol{\pi}})) \cup (\mathcal{K}(T_{\boldsymbol{\pi}}) \setminus \mathcal{S}_{\theta-\epsilon}) = \emptyset\right] \geq 1 - \delta, \tag{4.3}$$

for some user-specified *labeling error probability* $\delta \in (0, 1)$. Here, (4.3) enforces (4.2) by requiring that $\mathcal{K}(T_{\boldsymbol{\pi}})$ includes (almost) every one of the interesting cells and excludes (almost) every one of the uninteresting cells upon termination, with probability $1 - \delta$.

Additionally, the search must have:

1. low *priority labeling time* $L(\boldsymbol{\pi})$,

$$L(\boldsymbol{\pi}) = \inf\{t \leq T_{\boldsymbol{\pi}} : \mathbb{P}[\mathcal{S}_{\theta+\epsilon} \setminus \mathcal{K}(t) \neq \emptyset] \leq \delta]\}, \tag{4.4}$$

   i.e., the search identify (almost) every one of the interesting cells quickly. By definition, $L(\boldsymbol{\pi}) \leq T_{\boldsymbol{\pi}}$.

2. low *economic cost* upon termination, i.e., the search has low costs associated with movement and sensing,

$$E(\boldsymbol{\pi}) = \sum_{t=1}^{T_{\boldsymbol{\pi}}} \left(\underbrace{\ell(\mathbf{a}_t, \mathbf{a}_{t-1})}_{\text{movement cost}} + \underbrace{\beta d}_{\text{sensing cost}} \right) \tag{4.5}$$

   where $\mathbf{a}_t$ is the set of d grid cells being sampled at time t according to search policy $\boldsymbol{\pi}_{t-1}$, $\ell : \mathcal{G}^d \times \mathcal{G}^d \to \mathbb{R}$ is a metric on $\mathcal{G}^d$, and $\beta > 0$ is a known constant sensing cost for each agent. Consequently, βd is the sensing cost for the team at each time step.

We pursue probabilistic performance metrics in (4.3) and (4.4) due to the uncertainty in sensing and perception.

We now state the two problems tackled by this paper: Design a multi-agent search algorithm on $\mathcal{G}$ that simultaneously satisfies the criteria (4.3), (4.4), and (4.5).

Determine upper bounds on the time to terminate the search (4.3), the priority labeling time (4.4), and the economic cost (4.5) for the proposed solution to Problem 4.3.

4.4 Proposed Solution

Algorithm 3 describes the proposed solution for Problem 4.3. It augments label-then-move search with a thresholding bandit-based search (inspired from (Locatelli *et al.*, 2016)) to satisfy the criteria in (4.3), (4.4), and (4.5). In Algorithm 3, the keep set $\mathcal{K}(t) \subseteq \mathcal{G}$ and the reject set $\mathcal{R}(t) \subseteq \mathcal{G}$ are the sets of grid cells labeled as *interesting* and *uninteresting* respectively, at the time instant t.

Algorithm 3 runs in a loop until all grid cells in $\mathcal{G}$ are assigned to $\mathcal{K}(t)$ or $\mathcal{R}(t)$ (or both). Each loop starts with a toss of a biased coin with the bias set to the aggressiveness parameter, $\alpha \in (0, 1)$.

When the current toss of the biased coin returns heads, we use *upper confidence bounds* typical of bandit-based algorithms (Locatelli *et al.*, 2016; Mason *et al.*, 2020; Lattimore and Szepesvári, 2020) to sample the unlabeled cells "most likely" to be interesting. Specifically, we choose d distinct cells that achieve the highest values of

acquisition function $J : \mathcal{G} \times \mathbb{N} \to \mathbb{R} \cup \{\infty\}$ at time t,

$$J_{\boldsymbol{\pi}}(i,t) = \hat{\mu}_{i,\boldsymbol{\pi}}(t) + U_{i,\boldsymbol{\pi}}(t,\delta), \tag{4.6a}$$

$$\hat{\mu}_{i,\boldsymbol{\pi}}(t) = \frac{\sum_{h \in \mathcal{H}_i(t)} h}{|\mathcal{H}_i(t)|}, \tag{4.6b}$$

$$U_{i,\boldsymbol{\pi}}(t,\delta) = 2\sqrt{\frac{2\log(\log_2(2|\mathcal{H}_i(t)|)) + \log\left(12|\mathcal{G}|/\delta\right)}{2|\mathcal{H}_i(t)|}}, \tag{4.6c}$$

with $\hat{\mu}_{i,\boldsymbol{\pi}}(t) = U_{i,\boldsymbol{\pi}}(t,\delta) = \infty$, whenever $\mathcal{H}_i(t) = \emptyset$.

Otherwise, we minimize the movement cost ℓ in (4.5) to decide on the next location to sample. Since ℓ is a metric, a search agent continues to sample its current cell in the next iteration, if the current cell is unlabeled.

Algorithm 3 Multi-agent search under noisy observation

Input : Set of grid cells $\mathcal{G}$, number of agents $d \in \mathbb{N}$, threshold $\theta \in (0,1)$, tolerance $\epsilon > 0$, labeling error probability $\delta \in (0,1)$, aggressiveness param. $\alpha \in (0,1)$

Output : $\{\mathcal{K}(t)\}_{t\geq 1}$, a sequence of (keep) sets of grid cells

Initialize time counter $t \leftarrow 1$

while $\mathcal{R}(t) \cup \mathcal{K}(t) \neq \mathcal{G}$ **do**

  if current toss of α-biased coin returns heads **then**

    Define $\mathbf{a}_t$ by selecting d distinct grid cells that score the highest values in $J_{\boldsymbol{\pi}}$ (4.6)

  else

    Define $\mathbf{a}_t$ by assigning each agent to a distinct unlabeled cell that minimizes ℓ in (4.5);

  end if

  Deploy the agents to grid cells $\mathbf{a}_t$ and update the history $\mathcal{H}(t)$ based on collected noisy sensors;

  Update sets of labeled grid cells;

$$\mathcal{K}(t) \leftarrow \{i \in \mathcal{G} | \hat{\mu}_{i,\boldsymbol{\pi}}(t) - U_{i,\boldsymbol{\pi}}(t,\delta) \geq \theta - \epsilon\}, \tag{4.7a}$$

$$\mathcal{R}(t) \leftarrow \{i \in \mathcal{G} | \hat{\mu}_{i,\boldsymbol{\pi}}(t) + U_{i,\boldsymbol{\pi}}(t,\delta) \leq \theta + \epsilon\}. \tag{4.7b}$$

  Increment time counter $t \leftarrow t+1$

end while

return $\{\mathcal{K}(t)\}_{t\geq 1}$

Finally, we complete the loop by updating the sets $\mathcal{K}(t+1)$ and $\mathcal{R}(t+1)$ using (4.7) based on the data collected in the iteration t. Since $U_{i,\boldsymbol{\pi}}(t,\delta)$ is a non-increasing function of $|\mathcal{H}_i(t)|$, the sets $\mathcal{K}(t)$ and $\mathcal{R}(t)$ are monotonic in t. The definitions used in

(4.7) are motivated by the desire to obtain *anytime guarantees* for Algorithm 3.

Proposition 4.4.1 (ANYTIME ALGORITHM)**.** *The following holds for Algorithm 3 at any time $t \geq 1$ with probability of at least $1-\delta$: $\mathcal{K}(t) \subseteq \mathcal{S}_{\theta-\epsilon}$ and $\mathcal{R}(t) \subseteq \mathcal{S}^{\complement}_{\theta+\epsilon}$.*

We provide the proof of Proposition 4.4.1 in Appendix C.1. By Proposition 4.4.1, Algorithm 3 yields a correct-by-construction (albeit incomplete) labeling of the grid cells, even when it is terminated prematurely.

We conclude this section by noting that Algorithm 3 simplifies to a label-then-move search when $\alpha = 0$, and a thresholding bandit-based search when $\alpha = 1$.

4.5 Performance Analysis

We now focus on Problem 4.3, and study the performance of Algorithm 3. We show that Algorithm 3 has finite time termination guarantees, and admits high likelihood upper bounds to the incurred economic costs and priority labeling time. These bounds are a natural consequence of the bandit framework which yields non-asymptotic performance guarantees under minimal modeling assumptions.

We will use the following problem-specific parameters for each cell $i \in \mathcal{G}$,

$$\Delta_i = |\mu_i - \theta| + \epsilon, \text{ and } \Omega_i = \min_{j \in \mathcal{S}_{\theta+\epsilon}} |\mu_j - \mu_i|. \tag{4.8}$$

Informally, Δ_i signifies the separation of the mean μ_i from the threshold, while Ω_i signifies the separation of the mean μ_i from the set $\mathcal{S}_{\theta+\epsilon}$. We will state our results using parameters ϕ_i and γ_i for each cell $i \in \mathcal{G}$,

$$\phi_i = \frac{1}{\Delta_i^2} \log\left(\frac{|\mathcal{G}|}{\delta} \log\left(\frac{|\mathcal{G}|}{\Delta_i^4 \delta}\right)\right), \tag{4.9a}$$

$$\gamma_i = \frac{1}{\Omega_i^2} \log\left(\frac{|\mathcal{G}|}{\delta} \log\left(\frac{|\mathcal{G}|}{\Omega_i^4 \delta}\right)\right). \tag{4.9b}$$

Similar to the bandit literature (Lattimore and Szepesvári, 2020), we will show that Δ_i and Ω_i together characterize the difficulty of the search problem in Theorem 4.5.2.

Remark 4.5.1. *For any two scalar functions $f, g : \mathbb{R} \to \mathbb{R}$, we write $f = O(g)$ if there exists a constant $C > 0$ and a scalar $x_0 \in \mathbb{R}$ such that $f(x) \leq Cg(x)$ for every $x \geq x_0$.*

Theorem 4.5.2 (Upper bounds for Algorithm 3)**.** *Each one of the following statements hold for Algorithm 3 with probability $1 - \delta$:*

1. *Algorithm 3 terminates at $T_{\boldsymbol{\pi}}$ and satisfies the low labeling error criterion* (4.3) *with*

$$T_{\boldsymbol{\pi}} \leq \max_{i \in \mathcal{D}_\Delta} O\left(\phi_i\right) + \frac{1}{d} \sum\nolimits_{i \in \mathcal{D}_\Delta^{\complement}} O\left(\phi_i\right), \tag{4.10}$$

   where $\mathcal{D}_\Delta$ is the union of a grid cell with the smallest Δ_i with a set of $d-1$ grid cells with the largest Δ_i among all cells $i \in \mathcal{G}$.

2. *The priority labeling time* (4.4) *for Algorithm 3 is bounded from above as follows,*

$$\begin{aligned} L(\boldsymbol{\pi}) \leq & \max_{i \in \mathcal{D}_\Omega} O\left(\phi_i\right) + \frac{1}{d} \sum\nolimits_{i \in \mathcal{S}_{\theta+\epsilon} \setminus \mathcal{D}_\Omega} O\left(\phi_i\right) \\ & + \frac{1}{d} \sum_{i \in \mathcal{S}_{\theta+\epsilon}^{\complement} \setminus \mathcal{D}_\Omega} \min\left\{ O\left(\gamma_i\right) + \frac{4(1-\alpha)|\mathcal{G}|^2}{\alpha\delta}, O\left(\phi_i\right) \right\}, \end{aligned} \tag{4.11}$$

   where $\mathcal{D}_\Omega$ is a set of d grid cells characterized by Δ_i, Ω_i, and α.

3. *The economic cost* (4.5) *incurred by Algorithm 3 is bounded from above (with $M = \max_{\boldsymbol{a}, \boldsymbol{a}' \in \mathcal{G}^d} \ell(\boldsymbol{a}, \boldsymbol{a}')$),*

$$\begin{aligned} E(\boldsymbol{\pi}) \leq & \, O(|\mathcal{G}| - 1) + d \max_{i \in \mathcal{D}_\Delta} O\left((M\alpha + \beta)\phi_i\right) \\ & + \sum\nolimits_{i \in \mathcal{D}_\Delta^{\complement}} O\left((M\alpha + \beta)\phi_i\right). \end{aligned} \tag{4.12}$$

See Appendix C.2 for a sketch of the proof of Theorem 4.5.2. In Theorem 4.5.2, *big-O* notation hides constants factors which are independent of system parameters.

From (4.10), Algorithm 3 may take more time to terminate when Δ_i is small for at least one grid cell, i.e., μ_i is close to θ for some $i \in \mathcal{G}$. Additionally, the upper bound on the termination time $T_{\boldsymbol{\pi}}$ does not have a purely inverse-linear relationship the number of agents d, i.e., $T_{\boldsymbol{\pi}}$ is not upper bounded by an expression containing only $\frac{1}{d}\sum_{i\in\mathcal{G}} O(\phi_i)$. Instead, the upper bound in (4.10) has an additional term $\max_{i\in\mathcal{D}_\Delta} O(\phi_i)$ independent of d, which corresponds to the diminishing benefit of significantly increasing the number of agents.

We now analyze the role played by the aggressiveness parameter α in the performance of Algorithm 3. We observe that the upper bound on $T_{\boldsymbol{\pi}}$ is independent of α, consistent with the intuition that Algorithm 3 with $\alpha > 0$ and label-then-move search (Algorithm 3 with $\alpha = 0$) takes the same number of iterations for a search problem with identical Δ_i. This is because all grid cells must be labeled at the end, and both approaches rely on similar concentration inequalities to label a cell.

Recall that Algorithm 3 simplifies to label-then-move search for $\alpha = 0$, where the deployments of the agents are decided solely based on the associated movement costs. Consequently, as seen from the upper bounds, such an approach incurs a low economic cost $E(\boldsymbol{\pi})$, but may incur a high priority labeling time $L(\boldsymbol{\pi})$. On the other hand, setting $\alpha = 1$ simplifies Algorithm 3 to a pure bandit-based search that samples grid cells based on the maxima of the acquisition function $J_{\boldsymbol{\pi}}$ (4.6). Consequently, as also seen from the upper bounds, such an the approach will result in low priority labeling time $L(\boldsymbol{\pi})$, but high economic cost $E(\boldsymbol{\pi})$. Thus, by varying $\alpha \in (0, 1)$, the the method can achieve the desired trade-off between the priority labeling time and the economic costs of search.

4.6 Experiments

We use a numerical simulation to compare Algorithm 3 to three baselines, `AdaSearch` (Rolf *et al.*, 2021), a pure bandit-based search (Algorithm 3 with $\alpha = 1$), and label-then-move search (Algorithm 3 with $\alpha = 0$).

We setup the multi-agent search problem as follows: Consider a search environment of 10×10 grid cells with mean $\mu_i = 0.85$ for interesting cells $i \in \mathcal{G}$ and $\mu_j = 0.15$ for uninteresting cells $j \in \mathcal{G}$. We set the team size $d = 5$ with randomly chosen starting locations. We also set 10 randomly chosen grid cells as interesting. We set the tolerance $\epsilon = 10^{-3}$, labeling error probability $\delta = 10^{-3}$, and the threshold parameter $\theta = 0.5$. We set sensing costs $\beta = 0.01$, and define ℓ as the sum of the Manhattan distance between the agents' current and next locations.

We adapt `AdaSearch` (Rolf *et al.*, 2021) to solve the multi-agent search problem. Recall that `AdaSearch` adjusts the number of samples collected at each cell based on any valid data-driven confidence bounds. In our implementation of `AdaSearch`, we utilized the confidence bounds defined in (4.6). We also assumed that the agents follow identical raster paths and recompute the sample visitation counts upon completing a loop around the environment. Unlike Algorithm 3, `AdaSearch` additionally requires the total number of interesting cells to label the cells.

We analyze how the different search strategies label interesting cells to the keep set $\mathcal{K}(t)$ as time progresses in the algorithm. Figure 4.2 shows the performance of the algorithms on 100 randomly generated search problems. Based on our experiments, we recommend the choice of $\alpha = 0.2$ for the given choice of ℓ and β.

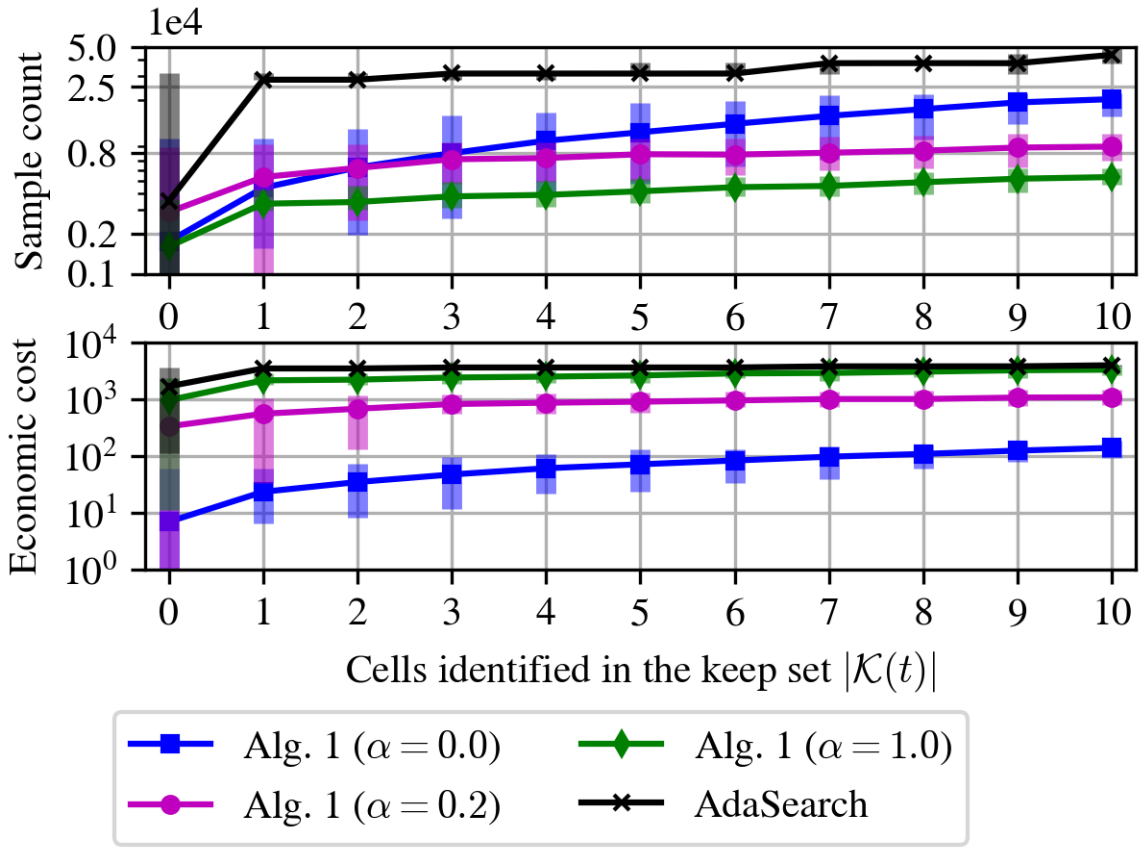


Figure 4.2: Priority labeling time for various search strategies with number of samples collected by the team over $\mathcal{G}$ (top), and incurred economic cost (bottom). The proposed solution (Algorithm 3 with $\alpha = 0.2$) achieves a good compromise as compared to other strategies — label-then-move search (Algorithm 3 with $\alpha = 0$), a pure bandit-based search (Algorithm 3 with $\alpha = 1$), and `AdaSearch` (Rolf *et al.*, 2021).

Priority labeling time (Fig. 4.2, top):

As expected, the proposed solution (Algorithm 3 with $\alpha = 0.2$) and a pure bandit-based search (Algorithm 3 with $\alpha = 1$) detects $\mathcal{S}_{\theta+\epsilon}$ with a smaller number of samples as compared to `AdaSearch` and label-then-move search (Algorithm 3 with $\alpha = 0$). The latter search strategies require a large amount of samples, possibly due to the pre-determined search pattern used by the agents.

Economic cost (Fig. 4.2, bottom):

The proposed solution and label-then-move search incur lower economic costs when compared to AdaSearch and pure bandit-based search. From (4.5), the economic cost is a linear combination of sampling cost and movement cost. Since β is small, the incurred costs are primarily driven by the movement costs, and the proposed solution and label-then-move search make the agents move relatively less when compared to other approaches. We expect the performance of the search strategies to be similar to Fig. 4.2, top, in applications where sensing is expensive and movement is cheap (higher β).

Impact of the team size d on the search (Fig. 4.3):

As expected, the number of samples needed per agent to characterize the keep set $\mathcal{K}(T_{\boldsymbol{\pi}})$ decreases with increasing team size d. However, the reduction in the samples needed per agent does not exhibit an inversely proportional relationship with the team size, as indicated by Theorem 4.5.2.

We conclude with a note that Algorithm 3 has minimal computational overhead. A non-optimized Python code took ≈ 0.3 milliseconds per iteration on a standard laptop.

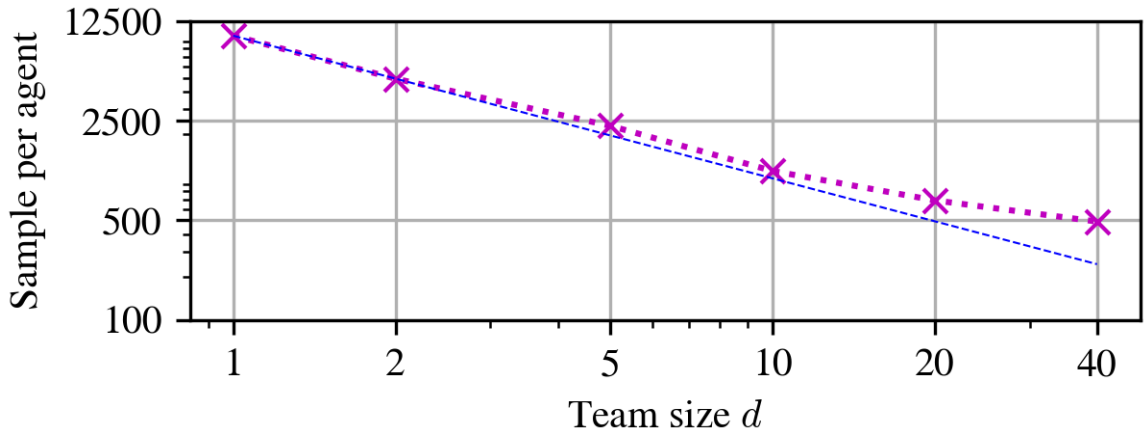


Figure 4.3: The median number of samples needed per agent to characterize the keep set $\mathcal{K}(T_{\boldsymbol{\pi}})$ (magenta line with crosses) in 100 randomly chosen search problems using the proposed solution (Algorithm 3 with $\alpha = 0.2$) decreases with increasing team size d. The blue line shows the trend needed to achieve an inversely proportional relationship between the samples needed per agent and the team size.

4.7 Conclusion & Future Work

We propose a data-driven, multi-agent search algorithm that accommodates noisy observations when searching for *all* interesting grid cells. We combined recent results from thresholding MABs with a standard label-then-move search to lower the time to identify interesting areas in the search environment and lower the costs incurred by the search agents during movement and sensing, while accommodating noisy observations.

The multi-agent search strategy proposed in this work has two drawbacks. First, it does not enforce the physical limitations on the mobile sensors are enforced as hard constraints during exploration. Second, it does not consider the effect of temporal changes in the search environment. Our future work will extend the proposed solution to address these drawbacks.

Chapter 5

PERIODIC BANDITS

5.1 Abstract

In traditional multi-armed bandits (MAB), a standard assumption is that the mean rewards are constant across each arm, a simplification that can be restrictive in nature. In many real-world settings, the rewards exhibit a periodic pattern on which traditional MAB algorithms would fail. This paper addresses the problem of regret minimization when the mean rewards change periodically. To this end, we propose an approach that utilizes the Ramanujan periodicity transform to estimate the support of the periods efficiently and, furthermore, use this information to minimize regret.

5.2 Introduction

Sequential decision-making under uncertainty is crucial in a wide variety of fields. Ideally, given ample time, one would exhaustively sample all available options before making decisions. However, in modern problems which present the decision maker with an enormous number of choices, such an approach is infeasible. The *Multi-armed bandits* (MAB) framework Slivkins *et al.* (2019) addresses this by efficiently identifying optimal options in minimal time. Central to MABs is the Exploration-exploitation dilemma: one must balance exploring unknown choices and exploiting the best-known option. Given its strong theoretical foundations and its efficacy in a wide range of domains like recommendation systems, clinical trials, and online advertising, this framework and its variants have received much attention in recent years Slivkins *et al.* (2019); Langford and Zhang (2007); Abbasi-yadkori *et al.* (2011).

A key limitation of this framework, however, is its traditional reliance on stationarity of the underlying "reward" distribution. Real-world applications often exhibit non-stationarity. Introducing non-stationary reward distributions complicates matters due to potential erratic patterns. Although there have been endeavors to address this (see, e.g., Besbes *et al.* (2019); Garivier and Moulines (2008); Slivkins and Upfal (2008)), formulating a universal learning policy for non-stationarity remains challenging.

In this paper, we focus on *Periodic Bandits*, a class of non-stationary bandits that are characterized by a periodic pattern in their rewards. Such periodicity is common in a range of real-world scenarios, such as cell-tower congestion, advertisement trends, and the behavior of electronic systems reliant on discharging power sources. Ignoring these patterns can result in highly suboptimal decisions Villamediana *et al.* (2019). Incorporating periodicity into multi-armed bandit algorithms enables one to make decisions that align more closely with the natural rhythms and temporal variations present in the problem domain.

Research such as Benedetto *et al.* (2020) has addressed seasonal reward shifts, while Re *et al.* (2021) leverages historical data for sudden changes. Other studies, like Zhou *et al.* (2021), focus on regime-switching rewards, while Chen *et al.* (2023b) considers rewards based on auto-regressive models. Cai *et al.* (2021) integrates periodicity in Gaussian process bandits. Our work aligns most closely with Chen *et al.* (2020), which combines Fourier analysis with a confidence-bound-based learning procedure to learn the periods and minimize the regret.

This paper proposes a tractable methodology for tacking the periodic bandit framework. To this end, we utilize the framework of Ramanujan Periodicity Transforms (RPT) to estimate the length of the period and identify the fundamental periods if the signal is a combination of two or more periodic signals. The authors in Tenneti and Vaidyanathan (2015); Vaidyanathan (2014) introduced the notion of RPT and showed

that one can utilize RPTs to estimate the underlying period of a periodic signal. In addition, the authors demonstrated that RPT-based methods are more robust in the presence of noise and showed the advantages of RPTs over the classical DFT-based techniques Tenneti and Vaidyanathan (2015). RPTs have been used in practice such as detecting periodicity in visually evoked potentials in brain-computer interfaces Saidi *et al.* (2019) and detecting the tandem DNA repeats Tenneti and Vaidyanathan (2016) and have shown promising results.

Contributions. The main contributions of this work are the following.

a) We propose an online learning algorithm called Bandit Tracking System via Ramanujan Periodic transform (BTS-RaP) for non-stationary environments with seasonal patterns and unknown periods.

b) We propose the use of RPT dictionaries to estimate the length of periods across different arms which are known to overcome the limitations of the DFT-based technique.

c) Using computer simulations we show that the BTS-RaP algorithm can achieve sublinear regret.

5.3 Ramanujan Periodicity Transforms

In this section, we briefly review the structure of the RPT dictionary, and their applicability to estimate the period of a periodic signal.

5.3.1 RPT Dictionaries

RPT dictionaries are constructed based on the properties of Ramanujan sums, defined as Ramanujan (1918)

$$c_p(n) = \sum_{\substack{k=1 \\ (k,p)=1}}^{p} \exp(j2\pi kn/p), \tag{5.1}$$

where (k, p) is the greatest common divisor of k and p. $\mathbf{c}_p$ indicates the vector form of $c_p(n)$, and $\mathbf{c}_p^{(i)}$ shows the circularly shifted version of $\mathbf{c}_p$ with step size i. For each value p construct a $p \times \phi(p)$ submatrix $\mathbf{C}_p$ as follows

$$\mathbf{C}_p = \begin{bmatrix} \mathbf{c}_p & \mathbf{c}_p^{(1)} & \dots & \mathbf{c}_p^{(\phi(p)-1)} \end{bmatrix}, \tag{5.2}$$

where $\phi(p)$ is the Euler totient function (the number of integers that are co-prime to p). The author in Vaidyanathan (2014) showed that the $\phi(p)$ columns in $\mathbf{C}_p$ are linearly independent. Thus, one can construct the RPT dictionary $\mathbf{K}$ in three consecutive steps.

i) Build all the submatrices $\mathbf{C}_p$ for every $p \in \mathbb{P}$, where $\mathbb{P} = \{1, 2, \dots, P_{\max}\}$ and $P_{\max}$ is the largest possible period in the signal.

ii) Build the $L \times \phi(p)$ submatrices $\mathbf{R}_p$, by periodically extending all the columns of $\mathbf{C}_p$ to length L.

iii) Concatenate the matrices $\mathbf{R}_p$ as

$$\mathbf{K} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 & \dots & \mathbf{R}_{P_{\max}} \end{bmatrix}. \tag{5.3}$$

Therefore, denoting $\phi(P_{\max}) = \sum_{p=1}^{P_{\max}} \phi(p)$, the size of the dictionary is $L \times \phi(P_{\max})$.

5.3.2 Period Estimation Using RPT Dictionary

Discrete periodic signals can be expressed using the RPT dictionary in a noise-free setup as:

$$\mathbf{y} = \mathbf{K}\mathbf{x} \tag{5.4}$$

where $\mathbf{y}$ is the vector form of the periodic signal with period p, $\mathbf{K}$ is the RPT dictionary introduced in section 5.3, and $\mathbf{x}$ is the sparse representation of the periodic signal under the RPT dictionary. Given a sufficiently long vector $\mathbf{y}$, vector $\mathbf{x}$ exhibits a sparse structure and its non-zero values correspond to the sub-matrices in $\mathbf{K}$, that

have periodic columns with periods q_i that are divisors of p, or $q_i|p$. Therefore, it is possible to estimate the period of a periodic signal by first recovering the sparse representation of the signal under the RPT dictionary. Then, the support set of the signal identifies the divisors of the underlying period of the signal. The support set of a sparse vector are a set of indices that contain the location of the non-zero values of the vector. Finally, the estimate of the period is equal to the least common multiplier (LCM) of the divisors from the recovered support set. One can recover the support of the sparse vector $\mathbf{x}$ using sparse recovery algorithms Baraniuk (2007). In this work, we adopt the proposed approach in Tenneti and Vaidyanathan (2015) and solve the following minimization program:

$$\min \|\mathbf{Dx}\|_2 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{Kx} \tag{5.5}$$

where, $\mathbf{D}$ is a diagonal penalty matrix with i-th entry on the diagonal being equal to p_i^2, where p_i is the period of the i-th column of the dictionary $\mathbf{K}$. An illustration of this method is presented in Fig. 5.1. In Part (a), we observe an incomplete segment of a signal with a period of 231, which has been affected by noise. This 231-period signal was constructed by combining three periodic signals with underlying periods of 3, 7, and 11. Following Tenneti and Vaidyanathan (2015), we compute the energy corresponding to each subvector in $\mathbf{x}$ as follows and plot the strength vs. period. Part (b) illustrates the strength vs. period after solving (5.5). The strength at each period p is defined as:

$$E(p) = \sum_{k=P+1}^{P+\phi(p)} |x(k)|^2, \quad P = \sum_{d=1}^{p-1} \phi(d). \tag{5.6}$$

Similarly, in part (c) the periodogram displays the strength of the different period components in the signal. It is evident that, RPT basis is robust towards estimating the fundamental periods of a given signal.

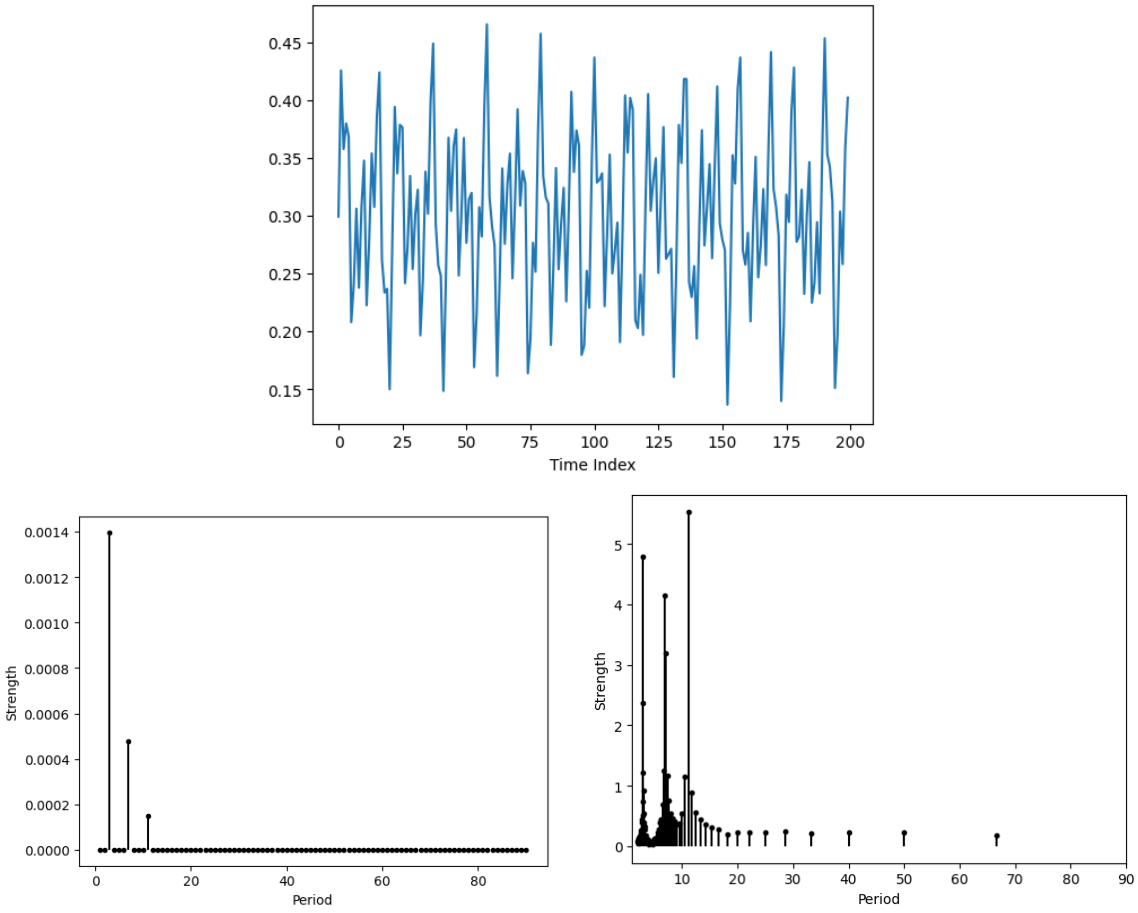


Figure 5.1: (a) A noisy period 231 time series signal with that was generated as sum of period 3, 7 and 11 signals. The strength vs period plot for the solutions of the convex problem (5.5) using (b) Ramanujan basis, and (c) DFT basis.

5.4 Problem Setup

Consider a multi-armed bandit setting with $\mathcal{K}$ being the set of all arms such that mean of each arm $i \in \mathcal{K}$ is represented by function $\mu_i : \mathbb{N} \to [a, b] \quad \forall i \in [K]$ such that $\mu_i[t + T_i] = \mu_i[t]$ for some unknown $T_i \in \mathbb{N}$. Throughout the paper, we sometimes refer $\mu_i[t]$ as $\mu_{t,i}$. At each round, the learner chooses an arm $a_t \in \mathcal{K}$ to sample and

observes a noisy reward

$$r_{t,i} = \mu_{t,i} + \eta_{t,i}\,,$$

where, $\{\eta_{i,t}\}_{i,t}$ are i.i.d. noise samples from a σ^2- sub-Gaussian distribution. The goal of the problem is to minimize the regret up to a known time horizon T defined as,

$$\mathcal{R}(T) = \sum_{t=0}^{T} \left(\max_{i \in \mathcal{K}} \mu_{t,i} - \mu_{t,a_t} \right)\,, \tag{5.7}$$

where the decision maker chooses an arm a_t at time step t. The aim of the work is to propose an algorithm to minimize regret as mentioned in (5.7). This can only be obtained if the decision maker chooses an arm that is optimal *for every time* t. If, at the current time instant, one of the arms is optimal, it is not necessary that the previously chosen arm will be optimal again at the next time step, which is well-suited for handling time-varying reward changes. The notion in (5.7) is different than the standard notion of regret Bubeck *et al.* (2015), which focuses on selecting the one optimal choice for every time-step t.

5.4.1 Baseline Method

Recently Chen *et al.* (2020) proposed for addressing periodic bandits a two-stage approach which provides a sub-linear regret that scales as $\mathcal{O}\left(\sqrt{T \sum_{i=1}^{n} T_i}\right)$, where T_i is the period of arm i. The authors first propose to use (DFT) to estimate the length of the periods T_i's. Since the mean of arm i returns to the same value every T_i step, the authors propose that for every arm i, the number of 'effective arms' is T_i (1 arm for every step until time reaches T_i). Therefore, we end up with $\hat{d} := \sum_k \hat{T}_k$ effective arms (unique mean rewards) to learn. In the second stage of the algorithm, the authors utilize the estimated number of effective arms to implement a UCB-based approach to minimize regret (5.7). We refer to this as MAB-UCB.

This approach suffers drawbacks due to the utilization of DFT as well as two distinct stages leading to sample inefficiency. We address this by using RPT and merging the two stages into one main algorithm thereby painting optimal sample efficiency.

5.5 Proposed Approach: BTS-RaP (Bandit Tracking System)

We first provide an overview of the linear bandits and then show how it connects to RPT-based reward representation.

5.5.1 Linear Bandits

Linear bandits Abbasi-yadkori *et al.* (2011) have emerged as a powerful and versatile tool in the field of bandit research literature. These algorithms are particularly well-suited for scenarios where the relationship between actions and rewards can be approximated linearly. Let the arm set be defined by the set $\mathcal{K}$. In a linear bandit setup, every arm is associated with a feature vector $\mathbb{R}^d$ such that $d < n$. On sampling the arm i at time t, the reward observed satisfies the relation $r_{t,i} = \langle \mathbf{a}_i, \boldsymbol{\theta}^* \rangle + \eta_t$, where $\boldsymbol{\theta}^* \in \mathbb{R}^d$ dented the unknown reward parameter, $\mathbf{a}_i \in \mathbb{R}^d$ denote the feature vector associated with arm i and η_t is the i.i.d. Gaussian noise realized from a σ^2-subgaussian distribution at time t.

Due to the low dimensional structure of the linear bandit problem, it has been proven, both theoretically and experimentally that the regret upper bound scales as $\mathcal{O}(\sqrt{dT})$ (sublinear in time T), where d is the feature dimensionality. Note that for the case of stationary linear bandit, the regret takes the form as defined below,

$$\mathcal{R}_{LB}(T) = \sum_{t=0}^{T} \left(\max_{k \in \mathcal{K}} \langle \mathbf{a}_k, \boldsymbol{\theta}^* \rangle - \langle \mathbf{a}_t, \boldsymbol{\theta}^* \rangle \right) \tag{5.8}$$

Taking a step further Wang *et al.* (2020) showcases that the regret upper bound can

be further tightened to nearly $\mathcal{O}(\sqrt{s_0 T})$, where $s_0 : \|\boldsymbol{\theta}^*\|_0 \leq s_0$ is the support of $\boldsymbol{\theta}^*$.

5.5.2 Connection to RPT Decomposition

Let $\mathbf{K}$ be the RPT dictionary and $\mathbf{x}_i$ be the corresponding sparse vector associated with the arm i with support set S_i. We can map our problem setup to linear bandits as follows:

i) Construct the block diagonal matrix $\mathbf{K}_{\mathcal{K}} = \text{diag}(\mathbf{K}, \mathbf{K}, \dots)$, where each $\mathbf{K} \in \mathbb{R}^{T \times \phi(P_{\max})}$ are blocks on the diagonal and constructed as per Equation (5.3). For an arm i, the arm features, at time t is $\mathbf{a}_{t,i} = \mathbf{K}_{\mathcal{K}}[i * T + t]$, which is the $(i * T + t)^{th}$ row of $\mathbf{K}_{\mathcal{K}}$.

ii) The unknown feature vector $\boldsymbol{\theta}^*$ is a vector stack of $\mathbf{x}_i$, where $\mathbf{x}_i$ is the true solution to the minimization problem (5.5) in the noiseless case. The reward is obtained as $r_{t,i} = \langle \mathbf{a}_{t,i}, \mathbf{x}_i \rangle$.The pseudocode of the proposed algorithm is provided in Algorithm 4 Following directly from Wang *et al.* (2020), we can provide the following theoretical backing to BTS-RaP:

Theorem 5.5.1. *Let $\mathbf{x}_i$ be the sparse representation of a periodic signal under the RPT dictionary with support set $\mathcal{S}_i$ that has $|\mathcal{S}_i|$ nonzero values, for all $\mathbf{x}_i \;\; i \in \mathcal{K}$, then the regret of Bandit Tracking System (BTS-RaP) is upper bounded by $\mathcal{O}(\sqrt{T \sum_{i \in \mathcal{K}} |\mathcal{S}_i|})$.*

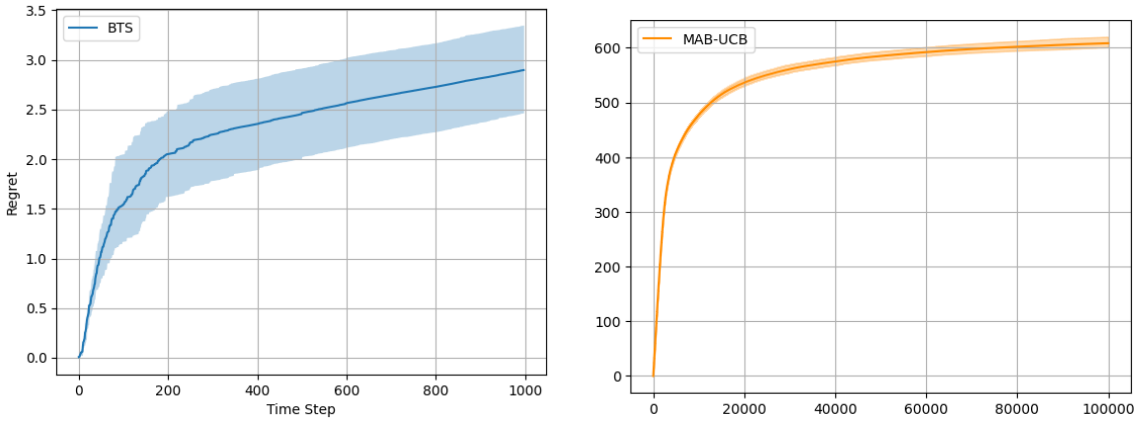


Figure 5.2: Regret $\mathcal{R}$ vs time t plots on two armed periodic bandits setting for (a) BTS-RaP and (b) MAB-UCB. Rewards of each arm is generated as per Equation (5.10).

Algorithm 4 Bandit Tracking System (BTS-RaP)

1: Given T, K arms, form $\mathbf{K} \in \mathbb{R}^{T \times \phi(P_{\max})}$ pull each arm once and form the observation vector $\boldsymbol{\mu}_i \ \forall i \in \mathcal{K}$

2: Create mini-dictionaries $\mathbf{K}_i$ for $i \in \mathcal{K}$ which will grow as arms get pulled. Initially each $\mathbf{K}_i$ is of size $1 \times \phi(P_{\max})$

3: Initialize support for each arm $\mathbf{x}_i$ for all $i \in \mathcal{K}$ $t = K+1 \ldots T$

4: Choose arm $a_t = \arg\max_{k \in \{1,\ldots,K\}} \langle \mathbf{K}[t], \mathbf{x}_i \rangle + \sqrt{2\alpha \ln t / n_{t-1,i}}$, where, $\mathbf{K}[t]$ is the t^{th} row of $\mathbf{K}$

5: Append the row $\mathbf{K}[t]$ to the mini-dictionary $\mathbf{K}_{a_t}$

6: Observe $r_{t,a_t} = \mu_{t,a_t} + \eta_t$ and append this observation to $\boldsymbol{\mu}_{a_t}$

7: Solve : $\min \|\mathbf{D}\mathbf{x}_{a_t}\|_2 \quad \text{s.t.} \quad \boldsymbol{\mu}_{a_t} = \mathbf{K}_{a_t}\mathbf{x}_{a_t}$ to return updated $\mathbf{x}_{a_t}$

8: $\{a_t\}_{t=K+1}^{T}$

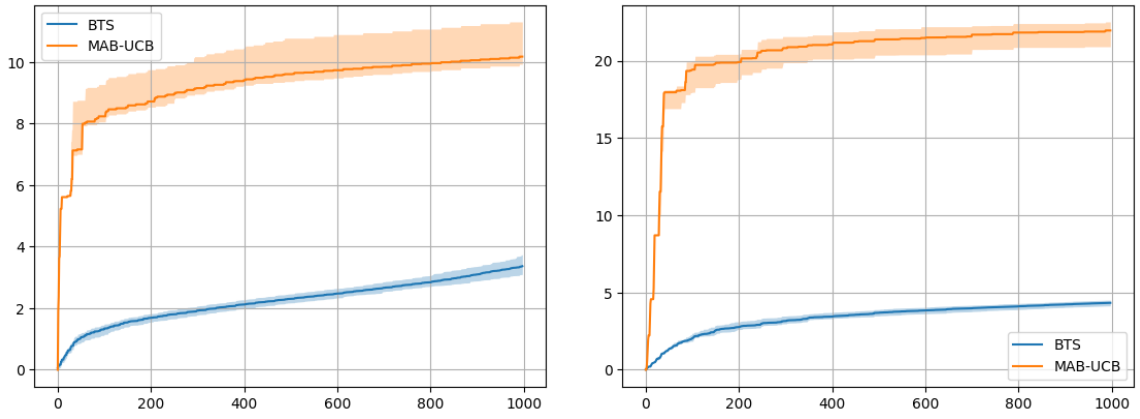


Figure 5.3: Regret $\mathcal{R}$ vs time t plots on two armed periodic bandits setting for MAB-UCB and BTS-RaP. Rewards of each arm is generated based on (5.9) with $\{p_1, p_2\}$ taking values (a) $\{7, 3\}$, (b) $\{9, 11\}$

5.6 Simulation Results

We consider a two-armed bandit setup with three different experiments. In the first and second experiments, we represent the means of the two arms by:

$$\mu_1(t) = c + \sin\left(\frac{2\pi t}{p_1}\right), \; \mu_2(t) = c + \sin\left(\frac{2\pi t}{p_2}\right), \tag{5.9}$$

where, $t = \{1, 2, \ldots, T\}$ and c is some positive scalar. For the first experiment the tuple $\{p_1, p_2\}$ take the values {9, 11} and for the second one it takes the values {7, 3}.

For the third experiment, we consider periodic mixtures to generate rewards as follows,

$$\mu_1(t) = c + \sum_{i=1}^{3} \sin\left(\frac{2\pi t}{p_i}\right), \; \mu_2(t) = c + \sin\left(\frac{2\pi t}{p}\right), \tag{5.10}$$

where, for the first arm the periods are $\{p_1, p_2, p_3\} = \{3, 7, 11\}$ and second arm period is $p = 9$. In the first two experiments (Figures 5.2(a), (b)), we see that our proposed algorithm BTS-RaP outperforms MAB-UCB. While plotting the regret of MAB-UCB we do not consider the stage one (estimation of period) cost. One advantage of using

RPT is that even if the period is large, we can still estimate it using RPT with fewer samples, sometimes, even when we have an incomplete period length signal as illustrated in Figure 1. While MAB-UCB does achieve sub-linear regret it does so at a very slow pace compared to BTS-RaP. This is because we are selecting an optimal arm from a set of $\sum_k T_k$ arms and not 2 as stated in the problem. This increases the complexity of the MAB problem and is reflected in the regret curve.

The real issue is revealed in the third experiment where one of the arms rewards is a combination of the sum of smaller periodic signals (7,3,11). Therefore, the resulting signal is a 231-length period signal. The second arm is a single period signal with period 9. So effectively, the MAB-UCB algorithm has 240 effective arms to select from. Whereas, BTS-RaP effectively learns non-zero coordinates of the support vector $\mathbf{x}$ associated with each arm. This vector as highlighted earlier is sparse and the regret scales with the sum ℓ_0-norm of this support vector. Therefore, as seen in Figure 5.3, BTS-RaP achieves minimum regret quickly and MAB-UCB has to run for a significantly longer time ($\sim 100\times$) to start learning the periodic pattern.

5.7 Conclusion

In this paper, we consider bandits that exhibit periodicity. We incorporated the periodic structure of the rewards and proposed an algorithm to minimize the regret. To this end, we utilized the newly introduced, Ramanujan-based periodicity estimation techniques to sequentially update the estimate of the periods of each arm, and subsequently select the best arm at each time step. Our results indicate that our RPT-based method dubbed BTS-RaP, can achieve sub-linear regret.

Abbasi-yadkori, Y., D. Pál and C. Szepesvári, "Improved algorithms for linear stochastic bandits", in "Advances in Neural Information Processing Systems", edited by J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira and K. Q. Weinberger, vol. 24 (Curran Associates, Inc., 2011), URL https://proceedings.neurips.cc/paper/2011/file/e1d5be1c7f2f456670de3d53c7b54f4a-Paper.pdf.

Ando, R. K. and T. Zhang, "Learning on graph with laplacian regularization", Advances in neural information processing systems **19**, 25 (2007).

Atsidakou, A., O. Papadigenopoulos, C. Caramanis, S. Sanghavi and S. Shakkottai, "Asymptotically-optimal Gaussian bandits with side observations", in "Proceedings of the 39th International Conference on Machine Learning", edited by K. Chaudhuri, S. Jegelka, L. Song, C. Szepesvari, G. Niu and S. Sabato, vol. 162 of *Proceedings of Machine Learning Research*, pp. 1057–1077 (PMLR, 2022), URL https://proceedings.mlr.press/v162/atsidakou22a.html.

Audibert, J.-Y., S. Bubeck and R. Munos, "Best arm identification in multi-armed bandits.", in "COLT", pp. 41–53 (2010).

Bajcsy, R., Y. Aloimonos and J. K. Tsotsos, "Revisiting active perception", A. Robots **42**, 2, 177–196 (2018).

Balan, R., "Reconstruction of signals from magnitudes of redundant representations: The complex case", Foundations of Computational Mathematics **16**, 3, 677–721 (2016).

Balan, R., P. Casazza and D. Edidin, "On signal reconstruction without phase", Applied and Computational Harmonic Analysis **20**, 3, 345–356 (2006).

Balan, R. and Y. Wang, "Invertibility and robustness of phaseless reconstruction", Applied and Computational Harmonic Analysis **38**, 3, 469–488 (2015).

Balan, R. and D. Zou, "Phase retrieval using lipschitz continuous maps", arXiv preprint arXiv:1403.2301 (2014).

Bandeira, A. S., J. Cahill, D. G. Mixon and A. A. Nelson, "Saving phase: Injectivity and stability for phase retrieval", Applied and Computational Harmonic Analysis **37**, 1, 106–125 (2014).

Banerjee, A., R. Ghods and J. Schneider, "Multi-agent active search using detection and location uncertainty", arXiv preprint arXiv:2203.04524 (2022).

Banerjee, T., G. Mukherjee and W. Sun, "Adaptive sparse estimation with side information", Journal of the American Statistical Association **115**, 532, 2053–2067 (2020).

Bapat, R. B. and S. Gupta, "Resistance distance in wheels and fans", Indian Journal of Pure and Applied Mathematics **41**, 1, 1–13 (2010).

Baraniuk, R. G., "Compressive sensing [lecture notes]", IEEE Signal Processing Magazine **24**, 4, 118–121 (2007).

Barvinok, A. I., "Problems of distance geometry and convex properties of quadratic maps", Discrete & Computational Geometry **13**, 2, 189–202 (1995).

Beck, A., "Convexity properties associated with nonconvex quadratic matrix functions and applications to quadratic programming", Journal of Optimization Theory and Applications **142**, 1, 1–29 (2009).

Beck, A. and Y. C. Eldar, "Strong duality in nonconvex quadratic optimization with two quadratic constraints", SIAM Journal on Optimization **17**, 3, 844–860 (2006).

Beck, A. and D. Pan, "A branch and bound algorithm for nonconvex quadratic optimization with ball and linear constraints", Journal of Global Optimization **69**, 2, 309–342 (2017).

Bendory, T., Y. C. Eldar and N. Boumal, "Non-convex phase retrieval from stft measurements", IEEE Transactions on Information Theory **64**, 1, 467–484 (2017).

Benedetto, G. D., V. Bellini and G. Zappella, "A linear bandit for seasonal environments", (2020).

Besbes, O., Y. Gur and A. Zeevi, "Optimal exploration-exploitation in a multi-armed-bandit problem with non-stationary rewards", Stochastic Systems **9**, 4, 319–337 (2019).

Best, G., J. Faigl and R. Fitch, "Online planning for multi-robot active perception with self-organising maps", A. Robots **42**, 4, 715–738 (2018).

Bhojanapalli, S., A. Kyrillidis and S. Sanghavi, "Dropping convexity for faster semi-definite optimization", in "Conference on Learning Theory", pp. 530–582 (2016a).

Bhojanapalli, S., B. Neyshabur and N. Srebro, "Global optimality of local search for low rank matrix recovery", in "Advances in Neural Information Processing Systems", pp. 3873–3881 (2016b).

Bodmann, B. G. and N. Hammen, "Algorithms and error bounds for noisy phase retrieval with low-redundancy frames", (2014).

Botelho-Andrade, S., P. G. Casazza, H. Van Nguyen and J. C. Tremain, "Phase retrieval versus phaseless reconstruction", Journal of Mathematical Analysis and Applications **436**, 1, 131–137, URL https://www.sciencedirect.com/science/article/pii/S0022247X15010847 (2016).

Bower, A., L. Jain and L. Balzano, "The landscape of non-convex quadratic feasibility", in "2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)", pp. 3974–3978 (IEEE, 2018).

Bubeck, S., O. Dekel, T. Koren and Y. Peres, "Bandit convex optimization:$\sqrt{T}$ regret in one dimension", in "Conference on Learning Theory", pp. 266–278 (PMLR, 2015).

Bubeck, S., R. Munos and G. Stoltz, "Pure exploration in multi-armed bandits problems", in "International conference on Algorithmic learning theory", pp. 23–37 (Springer, 2009).

Bubeck, S., R. Munos and G. Stoltz, "Pure exploration in finitely-armed and continuous-armed bandits", Theoretical Computer Science **412**, 19, 1832–1852 (2011).

Bullo, F., J. Cortés and S. Martinez, *Distributed control of robotic networks* (Princeton Univ. Press, 2009).

Cai, H., Z. Cen, L. Leng and R. Song, "Periodic-gp: Learning periodic world with gaussian process bandits", (2021).

Cai, T. T. and M. Yuan, "Adaptive covariance matrix estimation through block thresholding", The Annals of Statistics **40**, 4, 2014–2042 (2012).

Candes, E. J., "The restricted isometry property and its implications for compressed sensing", Comptes rendus mathematique **346**, 9-10, 589–592 (2008).

Candes, E. J., X. Li and M. Soltanolkotabi, "Phase retrieval via wirtinger flow: Theory and algorithms", IEEE Transactions on Information Theory **61**, 4, 1985–2007 (2015).

Candes, E. J., T. Strohmer and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming", Communications on Pure and Applied Mathematics **66**, 8, 1241–1274 (2013).

Cao, W., J. Li, Y. Tao and Z. Li, "On top-k selection in multi-armed bandits and hidden bipartite graphs", in "Advances in Neural Information Processing Systems", edited by C. Cortes, N. Lawrence, D. Lee, M. Sugiyama and R. Garnett, vol. 28 (Curran Associates, Inc., 2015), URL https://proceedings.neurips.cc/paper/2015/file/ab233b682ec355648e7891e66c54191b-Paper.pdf.

Chen, J., S. Huang, M. K. Ng and Z. Liu, "Solving quadratic systems with full-rank matrices using sparse or generative priors", (2023a).

Chen, N., C. Wang and L. Wang, "Learning and optimization with seasonal patterns", CoRR **abs/2005.08088**, URL https://arxiv.org/abs/2005.08088 (2020).

Chen, Q., N. Golrezaei and D. Bouneffouf, "Dynamic bandits with an auto-regressive temporal structure", (2023b).

Chen, S., T. Lin, I. King, M. R. Lyu and W. Chen, "Combinatorial pure exploration of multi-armed bandits.", in "NIPS", pp. 379–387 (2014).

Chen, Y., Y. Chi, J. Fan and C. Ma, "Gradient descent with random initialization: Fast global convergence for nonconvex phase retrieval", arXiv preprint arXiv:1803.07726 (2018).

Chung, F. R. and F. C. Graham, *Spectral graph theory*, no. 92 (American Mathematical Soc., 1997).

Dakic, T., *On the turnpike problem* (Simon Fraser University BC, Canada, 2000).

Dan, W., "The minimal measurement number for generalized conjugate phase retrieval", Science China Mathematics pp. 1–10 (2022).

Dasarathy, G., R. Nowak and X. Zhu, "S2: An efficient graph based active learning algorithm with application to nonparametric classification", in "Conference on Learning Theory", pp. 503–522 (PMLR, 2015).

Dasarathy, G., N. Rao and R. Baraniuk, "On computational and statistical tradeoffs in matrix completion with graph information", in "Signal Processing with Adaptive Sparse Structured Representations Workshop SPARS", (2017).

Davis, D., D. Drusvyatskiy and C. Paquette, "The nonsmooth landscape of phase retrieval", arXiv preprint arXiv:1711.03247 (2017).

Dines, L. L., "On the mapping of quadratic forms", Bulletin of the American Mathematical Society **47**, 6, 494–498 (1941).

Drémeau, A., A. Liutkus, D. Martina, O. Katz, C. Schülke, F. Krzakala, S. Gigan and L. Daudet, "Reference-less measurement of the transmission matrix of a highly scattering material using a dmd and phase retrieval techniques", Opt. Express **23**, 9, 11898–11911, URL https://opg.optica.org/oe/abstract.cfm?URI=oe-23-9-11898 (2015).

Drenth, J., *Principles of protein X-ray crystallography* (Springer Science & Business Media, 2007).

Drew, D. S., "Multi-agent systems for search and rescue applications", Current Robotics Rep. **2**, 189–200 (2021).

Du, B., K. Qian, H. Iqbal, C. Claudel and D. Sun, "Multi-robot dynamical source seeking in unknown environments", in "Intl Conf. Robotics and Autom.", pp. 9036–9042 (2021).

Du, S. S., J. D. Lee, H. Li, L. Wang and X. Zhai, "Gradient descent finds global minima of deep neural networks", arXiv preprint arXiv:1811.03804 (2018).

Duchi, J. C. and F. Ruan, "Solving (most) of a set of quadratic equalities: Composite optimization for robust phase retrieval", arXiv preprint arXiv:1705.02356 (2017).

Duxbury, P. M., L. Granlund, S. Gujarathi, P. Juhas and S. J. Billinge, "The unassigned distance geometry problem", Discrete Applied Mathematics **204**, 117–132 (2016).

Eldar, Y. C. and S. Mendelson, "Phase retrieval: Stability and recovery guarantees", Applied and Computational Harmonic Analysis **36**, 3, 473–494 (2014).

Even-Dar, E., S. Mannor, Y. Mansour and S. Mahadevan, "Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems.", Journal of machine learning research **7**, 6 (2006).

Gabillon, V., M. Ghavamzadeh and A. Lazaric, "Best arm identification: A unified approach to fixed budget and fixed confidence", in "NIPS-Twenty-Sixth Annual Conference on Neural Information Processing Systems", (2012).

Garivier, A. and E. Kaufmann, "Non-asymptotic sequential tests for overlapping hypotheses and application to near optimal arm identification in bandit models", (2019).

Garivier, A. and E. Moulines, "On upper-confidence bound policies for non-stationary bandit problems", arXiv preprint arXiv:0805.3415 (2008).

Ge, R., C. Jin and Y. Zheng, "No spurious local minima in nonconvex low rank problems: A unified geometric analysis", in "Proceedings of the 34th International Conference on Machine Learning-Volume 70", pp. 1233–1242 (JMLR. org, 2017).

Gentile, C., S. Li and G. Zappella, "Online clustering of bandits", (2014).

Ghods, R., A. Banerjee and J. Schneider, "Decentralized multi-agent active search for sparse signals", in "Proc. Conf. Uncertainty Artif. Intelli.", vol. 161, pp. 696–706 (2021).

Guo, D., Y. Wu, S. S. Shitz and S. Verdú, "Estimation in gaussian noise: Properties of the minimum mean-square error", IEEE Transactions on Information Theory **57**, 4, 2371–2385 (2011).

Guo, J., X. Cheng, G. Xu and H. Shen, "A structured approach to query recommendation with social annotation data", in "Proceedings of the 19th ACM international conference on Information and knowledge management", pp. 619–628 (2010).

Gupta, S., S. Chaudhari, G. Joshi and O. Yağan, "Multi-armed bandits with correlated arms", (2020).

Hager, W. W., "Updating the inverse of a matrix", SIAM Review **31**, 2, 221–239, URL http://www.jstor.org/stable/2030425 (1989).

Heinosaari, T., L. Mazzarella and M. M. Wolf, "Quantum tomography under prior information", Communications in Mathematical Physics **318**, 2, 355–374 (2013).

Huang, K., Y. C. Eldar and N. D. Sidiropoulos, "Phase retrieval from 1d fourier measurements: Convexity, uniqueness, and algorithms", IEEE Transactions on Signal Processing **64**, 23, 6105–6117 (2016).

Huang, K. and N. D. Sidiropoulos, "Consensus-admm for general quadratically constrained quadratic programming", Transactions on Signal Processing **64**, 20, 5297–5310 (2016).

Huang, M., Y. Rong, Y. Wang and Z. Xu, "Almost everywhere generalized phase retrieval", Applied and Computational Harmonic Analysis **50**, 16–33 (2021).

Huang, M. and Z. Xu, "The estimation performance of nonlinear least squares for phase retrieval", IEEE Transactions on Information Theory **66**, 12, 7967–7977 (2020).

Huang, S. and I. Dokmanić, "Reconstructing point sets from distance distributions", arXiv preprint arXiv:1804.02465 (2018).

Huang, S., S. Gupta and I. Dokmanić, "Solving complex quadratic systems with full-rank random matrices", arXiv preprint arXiv:1902.05612 (2019).

Ioannidis, V. N., X. Song, S. Manchanda, M. Li, X. Pan, D. Zheng, X. Ning, X. Zeng and G. Karypis, "Drkg - drug repurposing knowledge graph for covid-19", https://github.com/gnn4dr/DRKG/ (2020).

Iwen, M., A. Viswanathan and Y. Wang, "Fast phase retrieval from local correlation measurements", (2016).

Jaganathan, K., Y. C. Eldar and B. Hassibi, "Stft phase retrieval: Uniqueness guarantees and recovery algorithms", IEEE Journal of selected topics in signal processing **10**, 4, 770–781 (2016).

Jain, P., P. Kar *et al.*, "Non-convex optimization for machine learning", Foundations and Trends® in Machine Learning **10**, 3-4, 142–336 (2017).

Jamali, M. and M. Ester, "Trustwalker: a random walk model for combining trust-based and item-based recommendation", in "Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining", pp. 397–406 (2009).

Jamieson, K. and R. Nowak, "Best-arm identification algorithms for multi-armed bandits in the fixed confidence setting", in "2014 48th Annual Conference on Information Sciences and Systems (CISS)", pp. 1–6 (IEEE, 2014).

Ji, M. and J. Han, "A variance minimization criterion to active learning on graphs", in "Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics", edited by N. D. Lawrence and M. Girolami, vol. 22 of *Proceedings of Machine Learning Research*, pp. 556–564 (PMLR, La Palma, Canary Islands, 2012), URL http://proceedings.mlr.press/v22/ji12.html.

Jin, C., R. Ge, P. Netrapalli, S. M. Kakade and M. I. Jordan, "How to escape saddle points efficiently", in "Proceedings of the 34th International Conference on Machine Learning-Volume 70", pp. 1724–1732 (JMLR. org, 2017).

Jin, C., S. M. Kakade and P. Netrapalli, "Provable efficient online matrix completion via non-convex stochastic gradient descent", in "Advances in Neural Information Processing Systems", pp. 4520–4528 (2016).

Joshi, G., E. Soljanin and G. Wornell, "Efficient redundancy techniques for latency reduction in cloud systems", (2017).

Jun, K.-S., K. Jamieson, R. Nowak and X. Zhu, "Top arm identification in multi-armed bandits with batch arm pulls", in "Proc. Intl Conf. Artifi. Intelli. Stats.", vol. 51, pp. 139–148 (2016).

Kandasamy, K., G. Dasarathy, B. Poczos and J. Schneider, "The multi-fidelity multi-armed bandit", in "Advances in Neural Information Processing Systems", pp. 1777–1785 (2016).

Kapoutsis, A. C., S. A. Chatzichristofis and E. B. Kosmatopoulos, "DARP: Divide areas algorithm for optimal multi-robot coverage path planning", J. Intelli. Robotic Syst. **86**, 3, 663–680 (2017).

Karypis, G. and V. Kumar, "A fast and high quality multilevel scheme for partitioning irregular graphs", SIAM JOURNAL ON SCIENTIFIC COMPUTING **20**, 1, 359–392 (1998).

Kaufmann, E., O. Cappé and A. Garivier, "On the complexity of a/b testing", (2015).

Kaufmann, E., O. Cappé and A. Garivier, "On the complexity of best arm identification in multi-armed bandit models", (2016).

Klein, D. J. and M. Randić, "Resistance distance", Journal of mathematical chemistry **12**, 1, 81–95 (1993).

Kocák, T. and A. Garivier, "Best arm identification in spectral bandits", arXiv preprint arXiv:2005.09841 (2020).

Kolla, R. K., K. Jagannathan and A. Gopalan, "Collaborative learning of stochastic bandits over a social network", IEEE/ACM Transactions on Networking **26**, 4, 1782–1795 (2018).

Konar, A., "Non-convex quadratically constrained quadratic programming: Hidden convexity, scalable approximation and applications", (2017).

Konar, A. and N. D. Sidiropoulos, "Fast approximation algorithms for a class of non-convex qcqp problems using first-order methods", IEEE Transactions on Signal Processing **65**, 13, 3494–3509 (2017a).

Konar, A. and N. D. Sidiropoulos, "First-order methods for fast feasibility pursuit of non-convex qcqps", IEEE Transactions on Signal Processing **65**, 22, 5927–5941 (2017b).

Krause, A. and C. Guestrin, "Near-optimal observation selection using submodular functions", in "AAAI", vol. 7, pp. 1650–1654 (2007).

Kushnir, D. and L. Venturi, "Diffusion-based deep active learning", CoRR **abs/2003.10339**, URL https://arxiv.org/abs/2003.10339 (2020).

Langford, J. and T. Zhang, "The epoch-greedy algorithm for multi-armed bandits with side information", Advances in neural information processing systems **20** (2007).

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

Lee, J. D., I. Panageas, G. Piliouras, M. Simchowitz, M. I. Jordan and B. Recht, "First-order methods almost always avoid saddle points", arXiv preprint arXiv:1710.07406 (2017).

Lee, J. D., M. Simchowitz, M. I. Jordan and B. Recht, "Gradient descent converges to minimizers", arXiv preprint arXiv:1602.04915 (2016).

LeJeune, D., G. Dasarathy and R. Baraniuk, "Thresholding graph bandits with grapl", in "International Conference on Artificial Intelligence and Statistics", pp. 2476–2485 (PMLR, 2020).

Leskovec, J. and A. Krevl, "SNAP Datasets: Stanford large network dataset collection", http://snap.stanford.edu/data (2014).

Li, S., A. Karatzoglou and C. Gentile, "Collaborative filtering bandits", (2016).

Lipor, J. and G. Dasarathy, "Quantile search with time-varying search parameter", in "2018 52nd Asilomar Conference on Signals, Systems, and Computers", pp. 1016–1018 (IEEE, 2018).

Locatelli, A., M. Gutzeit and A. Carpentier, "An optimal algorithm for the thresholding bandit problem", in "International Conference on Machine Learning", pp. 1690–1698 (PMLR, 2016).

Luo, W., C. Nam, G. Kantor and K. Sycara, "Distributed environmental modeling and adaptive sampling for multi-robot sensor coverage", in "Proc. Intl. Conf. Auto. Agents Multi-Agent Syst.", pp. 1488–1496 (2019).

Ma, Y., R. Garnett and J. Schneider, "σ -optimality for active learning on gaussian random fields", in "Advances in Neural Information Processing Systems", edited by C. Burges, L. Bottou, M. Welling, Z. Ghahramani and K. Weinberger, vol. 26 (Curran Associates, Inc., 2013a), URL https://proceedings.neurips.cc/paper/2013/file/7810ccd41bf26faaa2c4e1f20db70a71-Paper.pdf.

Ma, Y., R. Garnett and J. G. Schneider, "σ-optimality for active learning on gaussian random fields.", in "NIPS", pp. 2751–2759 (2013b).

Ma, Y., T.-K. Huang and J. Schneider, "Active search and bandits on graphs using sigma-optimality", in "Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence", pp. 542–551 (2015a).

Ma, Y., T.-K. Huang and J. Schneider, "Active search and bandits on graphs using sigma-optimality", UAI'15, p. 542–551 (AUAI Press, Arlington, Virginia, USA, 2015b).

Mannor, S. and J. N. Tsitsiklis, "The sample complexity of exploration in the multi-armed bandit problem", **5**, 623–648 (2004).

Marchant, R. and F. Ramos, "Bayesian optimisation for intelligent environmental monitoring", in "IEEE Int'l Conf. Intelli. Robots Syst.", pp. 2242–2249 (2012).

Mason, B., L. Jain, A. Tripathy and R. Nowak, "Finding all ϵ-good arms in stochastic bandits", Adv. Neural Info. Process. Syst. **33**, 20707–20718 (2020).

Meckes, M. W., "Concentration of norms and eigenvalues of random matrices", Journal of Functional Analysis **211**, 2, 508–524 (2004).

Nordström, K., "Convexity of the inverse and moore–penrose inverse", Linear Algebra and its Applications **434**, 6, 1489–1512, URL https://www.sciencedirect.com/science/article/pii/S0024379510005999 (2011).

Park, J. and S. Boyd, "General heuristics for nonconvex quadratically constrained quadratic programming", arXiv preprint arXiv:1703.07870 (2017).

Pedarsani, R., D. Yin, K. Lee and K. Ramchandran, "Phasecode: Fast and efficient compressive phase retrieval based on sparse-graph codes", IEEE Transactions on Information Theory **63**, 6, 3663–3691 (2017).

Pólik, I. and T. Terlaky, "A survey of the s-lemma", SIAM review **49**, 3, 371–418 (2007).

Queralta, J., J. Taipalmaa, B. Pullinen, V. Sarker, T. Gia, H. Tenhunen, M. Gabbouj, J. Raitoharju and T. Westerlund, "Collaborative multi-robot search and rescue: Planning, coordination, perception, and active vision", IEEE Access **8**, 191617–191643 (2020).

Ramanujan, S., "On certain trigonometrical sums and their applications in the theory of numbers", Trans. Cambridge Philosoph. Soc **XXll**, 13, 259–276 (1918).

Rao, N., H.-F. Yu, P. Ravikumar and I. S. Dhillon, "Collaborative filtering with graph information: Consistency and scalable methods.", in "NIPS", vol. 2, p. 7 (Citeseer, 2015).

Re, G., F. Chiusano, F. Trovò, D. Carrera, G. Boracchi and M. Restelli, "Exploiting history data for nonstationary multi-armed bandit", in "Machine Learning and Knowledge Discovery in Databases. Research Track: European Conference, ECML PKDD 2021, Bilbao, Spain, September 13–17, 2021, Proceedings, Part I 21", pp. 51–66 (Springer, 2021).

Robbins, H., "Some aspects of the sequential design of experiments", Bulletin of the American Mathematical Society **58**, 5, 527–535 (1952).

Rolf, E., D. Fridovich-Keil, M. Simchowitz, B. Recht and C. Tomlin, "A successive-elimination approach to adaptive robotic source seeking", IEEE Tran. Robotics **37**, 1, 34–47 (2021).

Rozemberczki, B., C. Allen and R. Sarkar, "Multi-scale attributed node embedding", (2019).

Rozemberczki, B. and R. Sarkar, "Characteristic Functions on Graphs: Birds of a Feather, from Statistical Descriptors to Parametric Models", in "Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM '20)", p. 1325–1334 (ACM, 2020).

Sahni, S., "Computationally related problems", SIAM Journal on Computing **3**, 4, 262–279 (1974).

Saidi, P., A. Vosoughi and G. K. Atia, "Detection of brain stimuli using Ramanujan periodicity transforms", Journal of Neural Engineering **16**, 3, 036021 (2019).

Schwager, M., M. Vitus, S. Powers, D. Rus and C. J. Tomlin, "Robust adaptive coverage control for robotic sensor networks", IEEE Tran. Ctrl. Netw. Syst. **4**, 3, 462–476 (2015).

Shamir, O., "A variant of azuma's inequality for martingales with subgaussian tails", (2011).

Shunmugaraj, D. P., "Complex differentiation and cauchy reimann equations", URL http://home.iitk.ac.in/~psraj/mth102/lecture_notes/comp2.pdf (????).

Simon, H. A., "A behavioral model of rational choice", The quarterly journal of economics **69**, 1, 99–118 (1955).

Slivkins, A. and E. Upfal, "Adapting to a changing environment: the Brownian restless bandits.", in "COLT", pp. 343–354 (2008).

Slivkins, A. *et al.*, "Introduction to multi-armed bandits", Foundations and Trends® in Machine Learning **12**, 1-2, 1–286 (2019).

Soare, M., A. Lazaric and R. Munos, "Best-arm identification in linear bandits", (2014).

Spielman, D. A. and S.-H. Teng, "Spectral sparsification of graphs", (2010).

Sun, J., Q. Qu and J. Wright, "A geometric analysis of phase retrieval", Foundations of Computational Mathematics **18**, 5, 1131–1198 (2018).

Tan, Y. S. and R. Vershynin, "Phase retrieval via randomized kaczmarz: theoretical guarantees", Information and Inference: A Journal of the IMA (2017).

Tekin, C. and E. Turğay, "Multi-objective contextual multi-armed bandit with a dominant objective", IEEE Transactions on Signal Processing **66**, 14, 3799–3813 (2018).

Tenneti, S. V. and P. P. Vaidyanathan, "Nested periodic matrices and dictionaries: New signal representations for period estimation", IEEE Transactions on Signal Processing **63**, 14, 3736–3750 (2015).

Tenneti, S. V. and P. P. Vaidyanathan, "Detecting tandem repeats in DNA using Ramanujan filter bank", in "2016 IEEE International Symposium on Circuits and Systems (ISCAS)", pp. 21–24 (IEEE, 2016).

Thaker, P. K., G. Dasarathy and A. Nedić, "On the sample complexity and optimization landscape for quadratic feasibility problems", in "2020 IEEE International Symposium on Information Theory (ISIT)", pp. 1438–1443 (IEEE, 2020).

Vaidyanathan, P. P., "Ramanujan sums in the context of signal processing—part I : Fundamentals", IEEE Transactions on Signal Processing **62**, 16, 4145–4157 (2014).

Valko, M., R. Munos, B. Kveton and T. Kocák, "Spectral bandits for smooth graph functions", in "International Conference on Machine Learning", pp. 46–54 (PMLR, 2014).

Vershynin, R., "Introduction to the non-asymptotic analysis of random matrices", arXiv preprint arXiv:1011.3027 (2010).

Villamediana, J., I. Küster and N. Vila, "Destination engagement on facebook: Time and seasonality", Annals of Tourism Research **79**, 102747 (2019).

Vishnoi, N. K. *et al.*, "Lx= b", Foundations and Trends® in Theoretical Computer Science **8**, 1–2, 1–141 (2013).

Wang, D., J. Lipor and G. Dasarathy, "Distance-penalized active learning via markov decision processes", in "2019 IEEE Data Science Workshop (DSW)", pp. 155–159 (IEEE, 2019).

Wang, G., A. S. Zamzam, G. B. Giannakis and N. D. Sidiropoulos, "Power system state estimation via feasible point pursuit", in "Signal and Information Processing (GlobalSIP), 2016 IEEE Global Conference on", pp. 773–777 (IEEE, 2016).

Wang, Y., Y. Chen, E. X. Fang, Z. Wang and R. Li, "Nearly dimension-independent sparse linear bandit over small action spaces via best subset selection", (2020).

Wang, Y. and Z. Xu, "Generalized phase retrieval: measurement number, matrix recovery and beyond", Applied and Computational Harmonic Analysis (2017).

Wu, L., H.-F. Yu, N. Rao, J. Sharpnack and C.-J. Hsieh, "Graph dna: Deep neighborhood aware graph encoding for collaborative filtering", in "International Conference on Artificial Intelligence and Statistics", pp. 776–787 (PMLR, 2020).

Wu, Y., A. György and C. Szepesvári, "Online learning with gaussian payoffs and side observations", URL https://arxiv.org/abs/1510.08108 (2015).

Xiao, W. and I. Gutman, "Resistance distance and laplacian spectrum", Theoretical chemistry accounts **110**, 4, 284–289 (2003).

Yang, K., X. Dong and L. Toni, "Laplacian-regularized graph bandits: Algorithms and theoretical analysis", (2020).

Young, M. T., J. Hinkle, A. Ramanathan and R. Kannan, "Hyperspace: Distributed bayesian hyperparameter optimization", in "2018 30th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)", pp. 339–347 (IEEE, 2018).

Zhang, J., J. Katz-Samuels and R. D. Nowak, "GALAXY: graph-based active learning at the extreme", CoRR **abs/2202.01402**, URL https://arxiv.org/abs/2202.01402 (2022).

Zhou, X., Y. Xiong, N. Chen and X. Gao, "Regime switching bandits", Advances in Neural Information Processing Systems **34**, 4542–4554 (2021).

Zhou, Y., X. Chen and J. Li, "Optimal pac multiple arm identification with applications to crowdsourcing", in "International Conference on Machine Learning", pp. 217–225 (PMLR, 2014).

Zhu, X. J., "Semi-supervised learning literature survey", (2005).

APPENDIX A

GRAPH BANDITS

The appendix is organized as follows. Appendices A.1-A.3 and Appendix A.7 provide various supporting results and insights into our main theoretical results. Appendix A.4 and Appendix A.6 provide sample complexity guarantees for GRUB and ζ-GRUB respectively. Appendix A.5 states and proves necessary conditions on the sample complexity, and Appendix 2.8 presents a discussion on the incomparability of our graph bandits problem with that of linear bandits.

A.1 Parameter Estimation

At any time T, GRUB, along with the graph-side information, uses data gathered to estimate the mean $\hat{\boldsymbol{\mu}}_T$ in order to decide the sampling and elimination protocols. The following lemma gives the estimation routine used for GRUB.

Lemma A.1.1. *The closed form expression of $\hat{\boldsymbol{\mu}}_T$ is given by,*

$$\hat{\boldsymbol{\mu}}_T = \left(\sum_{t=1}^{T} \mathbf{e}_{\pi_t}\mathbf{e}_{\pi_t}^T + \rho L_G\right)^{-1} \left(\sum_{t=1}^{T} \mathbf{e}_{\pi_t} r_t^{\pi_t}\right) \tag{A.1}$$

Proof. Using the reward data $\{r_{t,\pi_t}\}_{t=1}^T$ gathered up-to time T and the sampling policy $\boldsymbol{\pi}_T$, the mean vector estimate $\hat{\boldsymbol{\mu}}_T$ is computed by solving the following laplacian-regularized least-square optimization schedule:

$$\hat{\boldsymbol{\mu}}_T = \underset{\boldsymbol{\mu}\in\mathbb{R}^n}{\arg\min} \quad \sum_{t=1}^{T} \left(\mu_{\pi_t} - r_{t,\pi_t}\right)^2 + \rho\langle \boldsymbol{\mu}, L_G\boldsymbol{\mu}\rangle \tag{A.2}$$

where $\rho > 0$ is a tunable penalty parameter. The above optimization problem can be equivalently written in the following quadratic form:

$$\hat{\boldsymbol{\mu}}_T = \underset{\boldsymbol{\mu}\in\mathbb{R}^n}{\arg\min} \left(\langle \boldsymbol{\mu}, V(\boldsymbol{\pi}_T, G)\boldsymbol{\mu}\rangle - 2\left\langle \boldsymbol{\mu}, \left(\sum_{t=1}^{T} \mathbf{e}_{\pi_t} r_{t,\pi_t}\right)\right\rangle + \sum_{t=1}^{T} r_{t,\pi_t}^2\right)$$

where $V(\boldsymbol{\pi}_T, G)$ denotes,

$$V(\boldsymbol{\pi}_T, G) = \sum_{t=1}^{T} \mathbf{e}_{\pi_t}\mathbf{e}_{\pi_t}^T + \rho L_G \tag{A.3}$$

In order to obtain $\hat{\boldsymbol{\mu}}_T$, we compute vanishing point of the gradient as follows,

$$\begin{aligned} &\left(\langle \boldsymbol{\mu}, V(\boldsymbol{\pi}_T, G)\boldsymbol{\mu}\rangle - 2\left\langle \boldsymbol{\mu}, \left(\sum_{t=1}^{T} \mathbf{e}_{\pi_t} r_{t,\pi_t}\right)\right\rangle + \sum_{t=1}^{T} r_{t,\pi_t}^2\right) |_{\boldsymbol{\mu}=\hat{\boldsymbol{\mu}}_T} = 0 \\ \Rightarrow \quad &\hat{\boldsymbol{\mu}}_T = V(\boldsymbol{\pi}_T, G)^{-1} \left(\sum_{t=1}^{T} \mathbf{e}_{\pi_t} r_t^{\pi_t}\right) \end{aligned} \tag{A.4}$$

□

The sampling policy in GRUB uses the mean estimates and their high probability confidence bounds to eliminate suboptimal arm. In the following lemma we compute the high probability confidence bounds on the estimates of the mean and introduces the idea of effective samples of each arm given the graph side information.

Lemma A.1.2. *For any $T > k(G)$ and $i \in [n]$, the following holds with probability no less than $1 - \frac{\delta}{w_i(\boldsymbol{\pi}_T)}$:*

$$|\hat{\mu}_T^i - \mu_i| \leq \sqrt{\frac{1}{t_{eff,i}}} \left(2\sigma \sqrt{14 \log\left(\frac{2w_i(\boldsymbol{\pi}_T)}{\delta}\right)} + \rho \|\boldsymbol{\mu}\|_G \right) \tag{A.5}$$

where $w_i(\boldsymbol{\pi}_T) = a_0 n t_{eff,i}^2$ for some constant $a_0 > 0$, $\hat{\mu}_T^i$ is the i-th coordinate of the estimate from A.1.1 and,

$$t_{eff,i} = \frac{1}{\left[\left(\sum_{t=1}^T \mathbf{e}_{\pi_t} \mathbf{e}_{\pi_t}^\top + \rho L_G \right)^{-1} \right]_{ii}}$$

Proof. Let the sequence of bounded variance noise and data gathered up-to time T be denoted by $\{\eta_t, r_{\pi_t,t}\}_{t=1}^T$. Let $S_T = \sum_{t=1}^T \eta_t \mathbf{e}_{\pi_t}$ and $N_T = \sum_{t=1}^T \mathbf{e}_{\pi_t} \mathbf{e}_{\pi_t}^T$. Using the closed form expression of $\hat{\boldsymbol{\mu}}_T$ from eq. A.1.1, the difference between the estimate and true value $\hat{\mu}_T^i - \mu_i$ can be obtained as follows:

$$\hat{\mu}_T^i - \mu_i = \langle \mathbf{e}_i, \hat{\boldsymbol{\mu}}_T - \boldsymbol{\mu} \rangle = \langle \mathbf{e}_i, V_T^{-1} S_T - \rho V_T^{-1} L_G \boldsymbol{\mu} \rangle$$

The deviation $\hat{\mu}_T^i - \mu_i$ can be upper-bounded as follows:

$$|\langle \mathbf{e}_i, \hat{\boldsymbol{\mu}}_T - \boldsymbol{\mu} \rangle| \leq |\langle \mathbf{e}_i, V_T^{-1} S_T \rangle| + |\langle \mathbf{e}_i, \rho V_T^{-1} L_G \boldsymbol{\mu} \rangle|$$

Further, in order to obtain the variance of the estimate $\hat{\boldsymbol{\mu}}_T$, we bound the deviation $|\mu_T^i - \mu_i|$ by separately bounding $|\langle \mathbf{e}_i, V_T^{-1} S_T \rangle|$ and $|\langle \mathbf{e}_i \rho V_T^{-1} L_G \boldsymbol{\mu} \rangle|$.

With regards to the first term $\langle \mathbf{e}_i, V_T^{-1} S_T \rangle$, note that

$$\begin{aligned} \langle \mathbf{e}_i, V_T^{-1} S_T \rangle &= \left\langle \mathbf{e}_i, V_T^{-1} \left(\sum_{t=1}^T \mathbf{e}_{\pi_t} \eta_t \right) \right\rangle \\ &= \sum_{t=1}^T \left\langle \mathbf{e}_i, V_T^{-1} \mathbf{e}_{\pi_t} \right\rangle \eta_t \end{aligned}$$

Using a variant of Azuma's inequality Shamir (2011); Valko *et al.* (2014), for any $\kappa > 0$ the following inequality holds,

$$\mathbb{P}\left(|\langle \mathbf{e}_i, V_T^{-1} S_T \rangle|^2 \leq \kappa^2 \right) \geq 1 - 2 \exp\left\{ - \frac{\kappa^2}{56\sigma^2 \sum_{t=1}^T \left(\langle \mathbf{e}_i, V_T^{-1} \mathbf{e}_{\pi_t} \rangle \right)^2} \right\} \tag{A.6}$$

Using the fact that $V_T \succ \left(\sum_{t=1}^{T} \mathbf{e}_{\pi_t}\mathbf{e}_{\pi_t}^T\right)$, we can further simplify the above bound using the following computation,

$$\begin{aligned}\sum_{t=1}^{T}\left(\left\langle \mathbf{e}_i, V_T^{-1}\mathbf{e}_{\pi_t}\right\rangle\right)^2 &= \left\langle V_T^{-1}\mathbf{e}_i, \left(\sum_{t=1}^{T}\mathbf{e}_{\pi_t}\mathbf{e}_{\pi_t}^T\right)V_T^{-1}\mathbf{e}_i\right\rangle \\ &\leq \langle \mathbf{e}_i, V_T^{-1}\mathbf{e}_i\rangle = [V_T^{-1}]_{ii} \end{aligned} \tag{A.7}$$

Substituting $\delta' = 2\exp\left\{-\frac{\kappa^2}{56\sigma^2\sum_{t=1}^{T}\left(\left\langle \mathbf{e}_i, V_T^{-1}\mathbf{e}_{\pi_t}\right\rangle\right)^2}\right\}$, we can finally conclude that given the historical data $\mathcal{F}_{T-1}$ till time $T-1$, following is true with probability $1-\delta'$,

$$|\langle \mathbf{e}_i, V_T^{-1}S_T\rangle|^2 \leq 56\sigma^2[V_T^{-1}]_{ii}\log\left(\frac{2}{\delta'}\right) \tag{A.8}$$

Second term $\langle \mathbf{e}_i, \rho V_T^{-1}L_G\boldsymbol{\mu}\rangle$ can be upperbounded using cauchy-schwartz inequality,

$$\begin{aligned}|\langle \mathbf{e}_i, \rho V_T^{-1}L_G\boldsymbol{\mu}\rangle| &= \rho\langle \mathbf{e}_i, L_G\boldsymbol{\mu}\rangle_{V_T^{-1}} \\ &\leq \rho\sqrt{\langle \mathbf{e}_i, V_T^{-1}\mathbf{e}_i\rangle}\sqrt{\langle L_G\boldsymbol{\mu}, V_T^{-1}L_G\boldsymbol{\mu}\rangle} \\ &\leq \rho\sqrt{[V_T^{-1}]_{ii}}\|\boldsymbol{\mu}\|_G \end{aligned} \tag{A.9}$$

Combining the upperbound (A.9), (A.8) and substituting $\delta' = \frac{\delta}{w(\boldsymbol{\pi}_T)}$ we get Lemma 2.3.2. Hence proved. □

A.2 Influence Factor

A key component in our characterization of the performance of GRUB is the *influence factor* for each arm; recall that for a given graph D, $C_i(D)$ denotes the connected component that contains i. The influence factor for each arm is defined as,

Definition A.2.1. *Let D be a graph on the vertex set $[n]$. For each $j \in [n]$, define* ***influence factor*** *$\mathfrak{I}(j, D)$ as:*

$$\mathfrak{I}(j,D) = \begin{cases} \min\limits_{i\in C_j(D), i\neq j}\{r_D(i,j)^{-1}\} & \text{if} \quad |C_j(D)| > 1 \\ 0 & \text{otherwise} \end{cases} \tag{A.10}$$

where, $r_D(i,j)$ is the resistance distance between arm i and j on graph D as in Definition 2.4.1.

Note that we refer the resistance distance without the parameter δ, as the value of resistance distance is independent of the value of δ. This happens due to the cancellation of δ factor in $R_{ii} + R_{jj} - R_{ji} - R_{ij}$. The influence factor can also be thought of as the minimum influence any arm i in the connected component of arm j has over the arm j

A.3 Effective Samples

Theorem A.3.1. *Let $\boldsymbol{\pi}_T$ indicate the sampling policy until time T. Let G be the given graph, $\mathfrak{I}(.,G)$ indicates the minimum influence factor for arms. Then effective samples can be lower bounded by,*

$$t_{eff,i} \geq t_i + \frac{1}{2}\lfloor \min\{\rho\mathfrak{I}(i,G), \sum_{j\in C(i)} t_j\}\rfloor \tag{A.11}$$

where t_i indicates the no. of samples of arm i and $\lfloor \cdot \rfloor$ indicates the floor.

Proof. Using Lemma A.7.5, we have the following bound on $[V_T^{-1}]_{ii}$,

$$[V(\boldsymbol{\pi}_T, G)^{-1}]_{ii} \leq \max\left\{\frac{1}{t_i + \frac{\rho\mathfrak{I}(i,G)}{2}}, \frac{1}{t_i + \frac{t_C - t_i}{2}}\right\} \tag{A.12}$$

where T is the total number of samples and t_C is all the samples from the connected component $C(i)$ apart from arm i. Thus rewriting the equation for $t_{\text{eff},i}$, we get,

$$t_{\text{eff},i} \geq t_i + \frac{1}{2}\min\{\rho\mathfrak{I}(i,G), \sum_{j\in C(i)} t_j\} \tag{A.13}$$

Hence proved. □

A.4 GRUB Sample Complexity

In order to compute the sample complexity for GRUB, we classify the arms into two categories: competitive and non-competitive. The split of arms into these two categories is not required for the algorithm, but provides tighter complexity bounds as will be observed in this appendix. The division of the arms is contingent on its suboptimality and the structure of the provided graph side information. A modified version of the Definition (2.4.3) of competitive set and non-competitive set is as follows:

Definition A.4.1. *Fix $\boldsymbol{\mu} \in \mathbb{R}^n$, graph D, regularization parameter ρ, confidence parameter δ, and smoothness parameter ϵ and noise variance σ. We define $\mathcal{H}$ to be the set of competitive arms and $\mathcal{N}$ to be the set of non-competitive arms as follows:*

$$\mathcal{H}(D,\boldsymbol{\mu},\delta,\rho,\epsilon) = \left\{ j \in [n] \middle| \Delta_i \leq 2\sqrt{\frac{2}{\rho\mathfrak{I}(i)}}\left(2\sigma\sqrt{14\log\left(\frac{2a_0 n\rho^2\mathfrak{I}(i)^2}{\delta}\right)} + \rho\epsilon\right)\right\},$$

$$\mathcal{N}(D,\boldsymbol{\mu},\delta,\rho,\epsilon) \triangleq [n] \setminus \mathcal{H}(D,\boldsymbol{\mu},\delta,\rho,\epsilon)$$

When the context is clear, we will use suppress the dependence on the parameters in Definition A.4.1.

Further, we derive an expression for the worst-case sample complexity by analysing the number of samples required to eliminate arms with different difficulty levels, i.e. arms in competitive set and non-competitive set. We first derive the sample complexity results for the case when graph G is connected and then extend it to disconnected graphs.

Lemma A.4.2. *Consider n-armed bandit problem with mean vector $\boldsymbol{\mu} \in \mathbb{R}^n$. Let G be a given connected similarity graph on the vertex set $[n]$, and further suppose that $\boldsymbol{\mu}$ is ϵ-smooth. Define*

$$T_{sufficient} \triangleq \sum_{i \in \mathcal{H}} \frac{1}{\Delta_i^2} \left[c_1 \log \frac{c_2}{\delta \Delta_i} + \frac{\rho \epsilon}{2} \right] + \max_{i \in \mathcal{N}} \left\{ \frac{2}{\Delta_i^2} \left[c_1 \log \frac{c_2}{\delta \Delta_i} + \frac{\rho \epsilon}{2} \right] \right\} \tag{A.14}$$

Then, with probability at least $1 - \delta$, GRUB: (a) terminates in no more than $T_{sufficient}$ rounds, and (b) returns the best arm $a^* = \arg\max_i \mu_i$.

Proof. With out loss of generality, assume that $a^* = 1$. Let $\{t_i\}_{i=1}^n$ denote the number of plays of each arm upto time T. By Lemma 2.3.2, we can state that,

$$\mathbb{P}\left(|\hat{\mu}_T^i - \mu_i| \geq \gamma_i(\boldsymbol{\pi}_T) \right) \leq \frac{2\delta}{a_0 n t_{\text{eff},i}^2} \tag{A.15}$$

where, $\gamma_i(\boldsymbol{\pi}_T) = \beta_i(\boldsymbol{\pi}_T) \sqrt{t_{\text{eff},i}^{-1}}$ and $\beta_i(\boldsymbol{\pi}_T) = \left(2\sigma \sqrt{14 \log \left(\frac{2 a_0 n t_{\text{eff},i}^2}{\delta} \right)} + \rho \|\boldsymbol{\mu}\|_G \right)$.

As is reflected in the elimination policy (2.4), at any time t, arm 1 can be mistakenly eliminated in GRUB only if $\hat{\mu}_t^i > \hat{\mu}_t^1 + \gamma_i(\boldsymbol{\pi}_t) + \gamma_1(\boldsymbol{\pi}_t)$. Let T_s be the stopping time of GRUB, then the total failure probability for GRUB can be upper-bounded as,

$$\mathbb{P}(\text{Failure}) \ \leq \sum_{t=2}^{T_s} \sum_{i=2}^{n} \mathbb{P}\left(\hat{\mu}_t^i \geq \hat{\mu}_t^1 + \gamma_i(\boldsymbol{\pi}_t) + \gamma_1(\boldsymbol{\pi}_t) \right)$$

Note that $\mathbb{P}\left(\hat{\mu}_t^i \geq \hat{\mu}_t^1 + \gamma_i(\boldsymbol{\pi}_t) + \gamma_1(\boldsymbol{\pi}_t) \right) \leq \left[\mathbb{P}\left(\hat{\mu}_t^i \geq \mu^i + \gamma_i(\boldsymbol{\pi}_t) \right) + \mathbb{P}\left(\hat{\mu}_t^1 \leq \mu^1 - \gamma_1(\boldsymbol{\pi}_t) \right) \right]$, provided that $\gamma_i(\boldsymbol{\pi}_t), \gamma_1(\boldsymbol{\pi}_t) \leq \frac{\Delta_i}{2}$. Hence the failure probability can be upperbounded as,

$$\mathbb{P}(\text{Failure}) \ \leq \sum_{i=2}^{n} \sum_{t=2}^{T_s} \left[\mathbb{P}\left(\hat{\mu}_t^i \geq \mu^i + \gamma_i(\boldsymbol{\pi}_t) \right) + \mathbb{P}\left(\hat{\mu}_t^1 \leq \mu^1 - \gamma_1(\boldsymbol{\pi}_t) \right) \right] \tag{A.16}$$

conditioned on $\gamma_i(\boldsymbol{\pi}_T), \gamma_1(\boldsymbol{\pi}_T) \leq \frac{\Delta_i}{2}$.

Let $a_0 \geq 4 \sum_{t=1}^{\infty} t_{\text{eff},i}^{-2}$, then from Lemma 2.3.2,

$$\begin{aligned} \mathbb{P}(\text{Failure}) \ &\leq \sum_{i=2}^{n} \sum_{t=2}^{T_s} \frac{2\delta}{a_0 n t_{\text{eff},i}^2} \\ &\leq \delta \end{aligned} \tag{A.17}$$

The finiteness of the infinite sum of ${t_{\text{eff},i}}^{-2}$ can be found in Lemma A.7.13.

Thus, in order to keep $\mathbb{P}(\text{Failure}) \leq \delta$, it is sufficient if, at the time of elimination of arm i, we have enough samples to ensure,

$$\begin{aligned} \gamma_i(\boldsymbol{\pi}_T) \ &\leq \frac{\Delta_i}{2} \\ \sqrt{\frac{1}{t_{\text{eff},i}}} \left(2\sigma \sqrt{14 \log \left(\frac{2 a_0 n t_{\text{eff},i}^2}{\delta} \right)} + \rho \epsilon \right) \ &\leq \frac{\Delta_i}{2} \end{aligned} \tag{A.18}$$

In the absence of graph information, equation (A.18) devolves to the same sufficiency condition for number of samples required for suboptimal arm elimination as Even-Dar *et al.* (2006), upto constant factor. Rewriting the above equation,

$$\frac{\log(a_i)}{a_i} \leq \sqrt{\frac{\delta}{d_1}}\frac{\Delta_i^2}{d_0} \tag{A.19}$$

where $d_0 = 64 \times 14\sigma^2, d_1 = 2na_0 e^{\frac{\rho^2\epsilon^2}{4\times 14\sigma^2}}$ and $a_i = \sqrt{\frac{d_1}{\delta}} t_{\text{eff},i}$. The following bound on a_i is sufficient to satisfy eq. (A.19),

$$a_i \geq 2\sqrt{\frac{d_1}{\delta}}\frac{d_0}{\Delta_i^2}\log\left(\sqrt{\frac{d_1}{\delta}}\frac{d_0}{\Delta_i^2}\right)$$

Resubstituting $t_{\text{eff},i}$, we obtain the sufficient number of plays required to eliminate arm i as,

$$t_{\text{eff},i} \geq \frac{c_1}{\Delta_i^2}\left[\log\left(\frac{c_2}{\delta^{\frac{1}{2}}\Delta_i^2}\right) + c_3\right] \tag{A.20}$$

where $c_1 = 2 \times 64 \times 14\sigma^2$, $c_2 = 64 \times 14\sigma^2\sqrt{2na_0}$ and $c_3 = \frac{\rho^2\epsilon^2}{8\times 14\sigma^2}$. In the further text we are suppressing the powers of δ, Δ_i within the log factor as it adds only a constant multiple to the lower bound.

The further part of the proof we use the following bound on $t_{\text{eff},\cdot}$ from Theorem A.3.1 as follows:

$$t_{\text{eff},i} \geq t_i + \frac{1}{2}\min\{\rho\mathfrak{I}(i), T - t_i\} \quad \forall i \in [n] \tag{A.21}$$

Hence a sufficiency condition for the GRUB to produce the best-arm with probability $1-\delta$ is given when both the following conditions are satisfied,

$$t_i + \frac{\rho\mathfrak{I}(i)}{2} \geq \frac{1}{\Delta_i^2}\left[c_1\log\left(\frac{c_2}{\delta\Delta_i}\right) + \frac{\rho\epsilon}{2}\right] \tag{A.22}$$

and,

$$T + t_i \geq T \geq \frac{2}{\Delta_i^2}\left[c_1\log\left(\frac{c_2}{\delta\Delta_i}\right) + \frac{\rho\epsilon}{2}\right] \tag{A.23}$$

From the Definition A.4.1 of competitive arms $\mathcal{H}$ and non-competitive arms $\mathcal{N}$, we have,

$$\mathcal{H} = \left\{j \in [n] \middle| \Delta_i \leq 2\sqrt{\frac{2}{\rho\mathfrak{I}(i)}}\left(2\sigma\sqrt{14\log\left(\frac{2a_0 n\rho^2\mathfrak{I}(i)^2}{\delta}\right)} + \rho\epsilon\right)\right\} \tag{A.24}$$

After the first $\max_{i\in\mathcal{N}}\left\{\frac{2}{\Delta_i^2}\left[c_1\log\frac{c_2}{\delta\Delta_i}+\frac{\rho\epsilon}{2}\right]\right\}$ samples, all arms in $\mathcal{N}$ are eliminated. Further, let k_1 be the index of the first arm to be eliminated (in $\mathcal{H}$) and $t^*_{k_1}$ be the number of samples of arm k_1 before getting eliminated then the total number of additional time steps played until the arm k_1 is eliminated is at most $|\mathcal{H}|t^*_{k_1}$. Let k_2 be the index of the next arm in $\mathcal{H}$ to be eliminated. The number of additional plays until the next arm is eliminated is given by $(|\mathcal{H}|-1)[t^*_{k_2}-t^*_{k_1}]$ and so on.

Summing up all the samples required to converge to the optimal arm is given by, (let $t^*_{k_0}=0$)

$$\sum_{h=1}^{|\mathcal{H}|}(|\mathcal{H}|-h))[t^*_{k_h}-t^*_{k_{h-1}}]=\sum_{h=1}^{|\mathcal{H}|-1}t^*_{k_h}=\sum_{i\in\mathcal{H}/1}t^*_i \tag{A.25}$$

Hence the final sample complexity can be computed as follows:

- Number of plays required for arms in $\mathcal{H}$:

$$\sum_{i\in\mathcal{H}/1}t^*_i\geq\sum_{i\in\mathcal{H}/1}\frac{1}{\Delta_i^2}\left[c_1\log\frac{c_2}{\delta\Delta_i}+\frac{\rho\epsilon}{2}\right] \tag{A.26}$$

- Number of plays required for all the arms in $\mathcal{N}:=[n]/\mathcal{H}$ to be eliminated:

$$T\geq\max_{i\in\mathcal{N}}\left\{\frac{2}{\Delta_i^2}\left[c_1\log\frac{c_2}{\delta\Delta_i}+\frac{\rho\epsilon}{2}\right]\right\} \tag{A.27}$$

Hence the final sample complexity can be given by,

$$T_{\text{sufficient}}\triangleq\max_{i\in\mathcal{N}}\left\{\frac{2}{\Delta_i^2}\left[c_1\log\frac{c_2}{\delta\Delta_i}+\frac{\rho\epsilon}{2}\right]\right\}+\sum_{i\in\mathcal{H}/1}\frac{1}{\Delta_i^2}\left[c_1\log\frac{c_2}{\delta\Delta_i}+\frac{\rho\epsilon}{2}\right] \tag{A.28}$$

Hence proved. □

We extend Lemma A.4.2 to the case when graph G has disconnected clusters.

Note: The following theorem stated in the thesis has a typographical error in the equation for $T_{\text{sufficient}}$ in place of arg min it is supposed to be min.

Theorem A.4.3. *Consider n-armed bandit problem with mean vector $\boldsymbol{\mu}\in\mathbb{R}^n$. Let $\mathcal{G}$ be the set of subgraphs of given similarity graph G on the vertex set $[n]$, and further suppose that $\boldsymbol{\mu}$ is ϵ-smooth. Define*

$$T_{sufficient}\triangleq\min_{D\in\mathcal{G}}\sum_{C\in\mathcal{C}_D}\left[\sum_{i\in C\cap\mathcal{H}_D}\frac{1}{\Delta_i^2}\left[c_1\log\frac{c_2}{\delta\Delta_i}+\frac{\rho\epsilon}{2}\right]+\max_{i\in C\cap\mathcal{N}_D}\left\{\frac{2}{\Delta_i^2}\left[c_1\log\frac{c_2}{\delta\Delta_i}+\frac{\rho\epsilon}{2}\right]\right\}\right] \tag{A.29}$$

where $\Delta_i=\mu^*-\mu_i$ for all suboptimal arms, $\mathcal{H}_D$ and $\mathcal{N}_D$ are as in Definition A.4.1, $\mathcal{C}_D$ is the set of connected components of a subgraph $D\in\mathcal{G}$ and c_1,c_2 are constants independent of system parameters. Then, with probability at least $1-\delta$, GRUB: (a) terminates in no more than $T_{sufficient}$ rounds, and (b) returns the best arm $a^*=\arg\max_i\mu_i$.

Proof. Let $\mathcal{C}_G$ denote the connected components of graph G. From Lemma A.4.2, the number of samples for each connected component $C \in \mathcal{C}_G$ can be given as,

$$T_{\text{sufficient}} = \left[\sum_{i \in C \cap \mathcal{H}} \frac{1}{\Delta_i^2} \left[c_1 \log \frac{c_2}{\delta \Delta_i} + \frac{\rho \epsilon}{2} \right] + \max_{i \in C \cap \mathcal{N}} \left\{ \frac{2}{\Delta_i^2} \left[c_1 \log \frac{c_2}{\delta \Delta_i} + \frac{\rho \epsilon}{2} \right] \right\} \right] \quad \text{(A.30)}$$

We can obtain the sample complexity for obtaining the best arm by summing it over all the components $C \in \mathcal{C}$, gives us the sample complexity for GRUB while considering graph G.

$$T_{\text{sufficient}} = \sum_{C \in \mathcal{C}_G} \left[\sum_{i \in C \cap \mathcal{H}} \frac{1}{\Delta_i^2} \left[c_1 \log \frac{c_2}{\delta \Delta_i} + \frac{\rho \epsilon}{2} \right] + \max_{i \in C \cap \mathcal{N}} \left\{ \frac{2}{\Delta_i^2} \left[c_1 \log \frac{c_2}{\delta \Delta_i} + \frac{\rho \epsilon}{2} \right] \right\} \right] \quad \text{(A.31)}$$

Any subgraph D of graph G satisfies,

$$\langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle \leq \epsilon \Rightarrow \langle \boldsymbol{\mu}, L_D \boldsymbol{\mu} \rangle \leq \epsilon \quad \text{(A.32)}$$

As seen in Definition A.4.1, the influence factor is instrumental in deciding the competitive and non-competitive sets, which further dictates the sample complexity bounds. Further, notice from Lemma A.7.8 that the influence factor $\mathfrak{I}(i, D)$ is not monotonic when considering subgraph D of graph G. Hence considering a subgraph of G could potentially increase the number of non-competitive arms and provide us with a tighter bound on the performance for GRUB.

Hence $T_{\text{sufficient}}$ in (A.30) can be made tighter by considering the minimum value over the entire set of subgraphs $\mathcal{G}$. □

We next derive sample complexity upper bounds for GRUB in certain illuminating special cases.

Corollary A.4.4 (Isolated clusters)**.** *Consider the setup as in Theorem 2.4.4 with the further restriction that G consists of a subgraph F such that optimal node is isolated and arms $[2, \ldots, n]$ are split in k clusters and $\Delta_i \geq 2\sqrt{\frac{2}{\rho \mathfrak{I}(i,F)}} \left(2\sigma \sqrt{14 \log \left(\frac{2a_0 n \rho^2 \mathfrak{I}(i,F)^2}{\delta} \right)} + \rho \epsilon \right)$, $\forall i \in [2, \ldots, n]$. Define*

$$T_{\textit{sufficient}} \triangleq \sum_{C \in \mathcal{C}_F / 1} \max_{j \in C} \frac{2}{\Delta_j^2} \left[c_1 \log \left(\frac{c_2}{\delta \Delta_i} \right) + \frac{\rho \epsilon}{2} \right] \quad \text{(A.33)}$$

Then, with probability at least $1 - \delta$, GRUB: (a) terminates in no more than $T_{\textit{sufficient}}$ rounds, and (b) returns the best arm $a^* = \arg\max_i \mu_i$.

Corollary A.4.4 shows that in scenarios where the arms are well clustered, the sample complexity of GRUB can scale with the number of clusters, a quantity that is typically significantly smaller than the total number of nodes in the graph.

Corollary A.4.5 (Star graph)**.** *Consider the setup as in Theorem 2.4.4 with the further restriction that G consists of a star subgraph with the central node as the optimal arm and $\Delta_i \leq 2\sqrt{\frac{2}{\rho\mathfrak{I}(i,F)}}\left(2\sigma\sqrt{14\log\left(\frac{2a_0 n\rho^2\mathfrak{I}(i,F)^2}{\delta}\right)}+\rho\epsilon\right),\ \forall i \in [2,\ldots,n]$. Define*

$$T_{sufficient} \triangleq \sum_{i=2}^{n}\frac{1}{\Delta_i^2}\left[c_1\log\left(\frac{c_2}{\delta\Delta_i}\right)+\frac{\rho\epsilon}{2}\right] \tag{A.34}$$

Then, with probability at least $1-\delta$, GRUB: (a) terminates in no more than $T_{sufficient}$ rounds, and (b) returns the best arm $a^* = \arg\max_i \mu_i$.

In Corollary A.4.5, $T_{\text{sufficient}}$ is the same sample complexity as vanilla best arm identification, upto constant factors which is due to the fact that pulling one of the spoke arms does not yield much information about the other spoke arms, and this is the exact situation in the standard pure exploration setting.

A.5 Lower Bounds

In this section we give a lower bound on the sample complexity for any δ-PAC to return the best arm for a n armed bandit problem along with graph side information.

Theorem A.5.1. *Given an n-armed bandit model with associated mean vector $\boldsymbol{\mu} \in \mathbb{R}^n$ and similarity graph G smooth on $\boldsymbol{\mu}$, i.e. $\langle\boldsymbol{\mu}, L_G\boldsymbol{\mu}\rangle \leq \epsilon$, for any $0 < \epsilon < \epsilon_0$. Let $G = ([n], E)$ be the graph with only k isolated cliques and w.l.o.g let arm 1 be the optimal arm. Then define*

$$T_{necessary} = \sum_{C\in\mathcal{C}_G/C^*}\min_{j\in C}\left\{\frac{4\sigma^2\log 5}{(\Delta_j-\sqrt{\epsilon})^2}\right\}+\sum_{j\in C^*/1}\frac{4\sigma^2\log 5}{\Delta_j^2} \tag{A.35}$$

where C^* is the clique with the optimal arm and $\epsilon_0 := \min_{i\in[n]/1, j\in C(i)}\left[\Delta_j\left[1-\frac{\Delta_i}{\sqrt{\Delta_i^2+\Delta_j^2}}\right]\right]^2$. Then any δ-PAC algorithm will need at-least $T_{necessary}$ steps to terminate, provided $\delta \leq 0.1$.

Proof. We prove the theorem in two steps: Firstly, we construct the sample complexity lower bound for the similarity graph with the isolated optimal arm and a clique of rest of the sub-optimal arms, followed by step 2 the sample complexity lower bound for a graph with single cluster

Step 1:

Consider a $n+1$ armed bandit problem with mean vector $\boldsymbol{\mu} \in \mathbb{R}^{n+1}$ and similarity graph M with an isolated optimal arm (arm 1) and n-clique cluster of suboptimal arms, satisfying the condition for smoothness of rewards over the graph,i.e., $\langle\boldsymbol{\mu}, L_M\boldsymbol{\mu}\rangle \leq \epsilon$. Then the following holds

$$\max_{i\neq 1}\mu_i \leq \min_{j\neq 1}\{\mu_j+\sqrt{\epsilon}\} \tag{A.36}$$

Assume that ordering of mean in n-clique of suboptimal arms is known. From (Kaufmann *et al.*, 2015), there exists a δ-PAC algorithm, for $\delta \leq 0.1$, which can successful identify the best arm for the subproblem with just the optimal arm and arm with the maximum mean in the n-clique cluster, i.e. $j' = \arg\max_{j \neq 1} \mu_j$ with the total number of samples given by,

$$T \geq \frac{4 \log 5\sigma^2}{\Delta_{j'}^2} \tag{A.37}$$

Now consider the case where the ordering of the mean in n-clique is unknown. In order to remove all the suboptimal arms provided $\epsilon \leq \min_{j \neq 1} \Delta_j^2$ and (A.36) holds, it is suffices to be able to distinguish between the optimal arm and a hypothetical suboptimal arm with mean $\mu_j + \sqrt{\epsilon}$ where j is any arm from suboptimal n-clique, and the minimum number of samples required by any δ-PAC algorithm to successfully identify the best arm with $\delta \leq 0.1$ is given by,

$$T \geq \frac{4 \log 5\sigma^2}{(\Delta_j - \sqrt{\epsilon})^2} \tag{A.38}$$

The best performance in terms of sample complexity out of all the random choice of arm from the suboptimal n-clique cluster is,

$$T \geq \min_{j \neq 1} \left\{ \frac{4 \log 5\sigma^2}{(\Delta_j - \sqrt{\epsilon})^2} \right\} \tag{A.39}$$

Given $\epsilon_0 := \min_{i \in [n]/1, j \in C(i)} \left[\Delta_j \left[1 - \frac{\Delta_i}{\sqrt{\Delta_i^2 + \Delta_j^2}} \right] \right]^2$ and $\epsilon < \epsilon_0$, it can be verified that for any arm $i, j \neq 1$,

$$\min_{j \neq 1} \frac{4 \log 5\sigma^2}{(\Delta_j - \sqrt{\epsilon})^2} < \frac{4 \log 5\sigma^2}{\Delta_i^2} + \frac{4 \log 5\sigma^2}{\Delta_j^2} \tag{A.40}$$

where the left hand side corresponds to the sample complexity lower bound of removing the suboptimal arms i, j with the graph side information and the right hand side corresponds to the same without graph side information.

Hence it can be inferred that it is inefficient to remove the arms individually (disregarding the graph information).

Step 2 :

Consider a $n+1$ armed bandit problem with mean vector $\boldsymbol{\mu} \in \mathbb{R}^{n+1}$ with a given similarity graph N such that $\langle \boldsymbol{\mu}, L_N \boldsymbol{\mu} \rangle \leq \epsilon$. Let all the suboptimal arms be connected to the optimal arm.

Here we show by an adversarial example that it is not possible to have a lower bound on the sample complexity which scales better than,

$$T \geq \sum_{j \neq 1} \frac{4 \log 5\sigma^2}{\Delta_j^2} \tag{A.41}$$

There exists a δ-PAC algorithm which can determine that arms $j = 3, \ldots, n$ are suboptimal after $T \geq \sum_{j \neq 1,2} \frac{1}{\Delta_j^2}$ samples. From the smoothness of rewards on the similarity graph N we know that,

$$-\sqrt{\epsilon} \leq \mu_1 - \mu_j \leq \sqrt{\epsilon} \quad \forall j \in [2, 3, \ldots, n] \tag{A.42}$$

This information does not help us identify or even reduce the number of samples required to identify optimal arm between arm 1 and arm 2. Thus no δ-PAC algorithm, $\delta \leq 0.1$, can determine the optimal arm from arm 1 and arm 2 without an additional $\frac{4 \log 5 \sigma^2}{\Delta_2^2}$ samples for determining the best arm.

Using above two steps, we construct the proof for lower bound as follows:

Now consider the graph side information as defined in the theorem, and let $\mathcal{C}_G$ denote the set of connected components of graph G and $C^* \in \mathcal{C}_G$ be the component containing the optimal arm. Finding the best arm in this setup requires elimination of the suboptimal arms with in the connected component containing optimal arm $j \in C^*$ and elimination of the other connected components with suboptimal arms $j \in \mathcal{C}_G / C^*$. Hence, the sample complexity lower bounds (Kaufmann *et al.*, 2015, 2016) for any δ-PAC algorithm with $\delta \leq 0.1$ to eliminate these arms using the tools developed in step 1 and step 2, is given by

$$T \geq \sum_{j \in C^*/1} \frac{4\sigma^2 \log 5}{\Delta_j^2} + \sum_{C \in \mathcal{C}_G / C^*} \min_{j \in C} \left\{ \frac{4\sigma^2 \log 5}{(\Delta_j - \sqrt{\epsilon})^2} \right\} \tag{A.43}$$

□

A.6 ζ-GRUB Sample Complexity Proof

Definition A.6.1. *Fix $\boldsymbol{\mu} \in \mathbb{R}^n$, graph D, confidence parameter δ, noise variance σ, and relaxation parameter ζ. We define $\mathcal{H}$ to be the set of competitive arms and $\mathcal{N}$ to be the set of non-competitive arms for ζ-GRUB as follows:*

$$\mathcal{H}(D, \boldsymbol{\mu}, \delta, \zeta) = \left\{ j \in [n] \middle| \Delta_i^\zeta \leq 2\sqrt{\frac{2}{\rho \mathfrak{I}(i)}} \left(2\sigma \sqrt{14 \log \left(\frac{2 a_0 n \rho^2 \mathfrak{I}(i)^2}{\delta} \right)} + \rho\epsilon \right) \right\},$$

$$\mathcal{N}(D, \boldsymbol{\mu}, \delta, \zeta) \triangleq [n] \setminus \mathcal{H}(D, \boldsymbol{\mu}, \delta, \zeta)$$

where $\Delta_i^\zeta \triangleq \max\{\Delta_i, \zeta\}$.

Lemma A.6.2. *Consider n-armed bandit problem with mean vector $\boldsymbol{\mu} \in \mathbb{R}^n$. Let G be a given connected similarity graph on the vertex set $[n]$, and further suppose that $\boldsymbol{\mu}$ is ϵ-smooth. Define*

$$T_{sufficient} \triangleq \sum_{i \in \mathcal{H}} \frac{1}{(\Delta_i^\zeta)^2} \left[c_1 \log \frac{c_2}{\delta \Delta_i^\zeta} + \frac{\rho\epsilon}{2} \right] + \max_{i \in \mathcal{N}} \left\{ \frac{2}{(\Delta_i^\zeta)^2} \left[c_1 \log \frac{c_2}{\delta \Delta_i^\zeta} + \frac{\rho\epsilon}{2} \right] \right\} \tag{A.44}$$

where $\Delta_i^\zeta \triangleq \max\{\Delta_i, \zeta\}$. Then, with probability at least $1 - \delta$, GRUB: (a) terminates in no more than $T_{sufficient}$ rounds, and (b) returns a ζ-best arm

Proof. With out loss of generality, assume that $a^* = 1$. Let $\{t_i\}_{i=1}^n$ denote the number of plays of each arm upto time T. By Lemma 2.3.2, we can state that,

$$\mathbb{P}\left(|\hat{\mu}_T^i - \mu_i| \geq \gamma_i(\boldsymbol{\pi}_T)\right) \leq \frac{2\delta}{a_0 n t_{\text{eff},i}^2} \tag{A.45}$$

where, $\gamma_i(\boldsymbol{\pi}_T) = \beta_i(\boldsymbol{\pi}_T)\sqrt{t_{\text{eff},i}^{-1}}$ and $\beta_i(\boldsymbol{\pi}_T) = \left(2\sigma\sqrt{14\log\left(\frac{2a_0 n t_{\text{eff},i}^2}{\delta}\right)} + \rho\|\boldsymbol{\mu}\|_G\right)$.

As is reflected in the elimination policy (2.4), at any time t, arm 1 can be mistakenly eliminated in GRUB only if $\hat{\mu}_t^i > \hat{\mu}_t^1 + \gamma_i(\boldsymbol{\pi}_t) + \gamma_1(\boldsymbol{\pi}_t)$. Let T_s be the stopping time of GRUB, then the total failure probability for GRUB can be upper-bounded as,

$$\mathbb{P}(\text{Failure}) \ \leq \sum_{t=2}^{T_s}\sum_{i=2}^{n}\mathbb{P}\left(\hat{\mu}_t^i \geq \hat{\mu}_t^1 + \gamma_i(\boldsymbol{\pi}_t) + \gamma_1(\boldsymbol{\pi}_t)\right)$$

Note that $\mathbb{P}\left(\hat{\mu}_t^i \geq \hat{\mu}_t^1 + \gamma_i(\boldsymbol{\pi}_t) + \gamma_1(\boldsymbol{\pi}_t)\right) \leq \left[\mathbb{P}\left(\hat{\mu}_t^i \geq \mu^i + \gamma_i(\boldsymbol{\pi}_t)\right) + \mathbb{P}\left(\hat{\mu}_t^1 \leq \mu^1 - \gamma_1(\boldsymbol{\pi}_t)\right)\right]$, provided that $\gamma_i(\boldsymbol{\pi}_t), \gamma_1(\boldsymbol{\pi}_t) \leq \frac{\Delta_i^\zeta}{2}$. Hence the failure probability can be upperbounded as,

$$\mathbb{P}(\text{Failure}) \ \leq \sum_{i=2}^{n}\sum_{t=2}^{T_s}\left[\mathbb{P}\left(\hat{\mu}_t^i \geq \mu^i + \gamma_i(\boldsymbol{\pi}_t)\right) + \mathbb{P}\left(\hat{\mu}_t^1 \leq \mu^1 - \gamma_1(\boldsymbol{\pi}_t)\right)\right] \tag{A.46}$$

conditioned on $\gamma_i(\boldsymbol{\pi}_T), \gamma_1(\boldsymbol{\pi}_T) \leq \frac{\Delta_i^\zeta}{2}$.

Let $a_0 \geq 4\sum_{t=1}^{\infty} t_{\text{eff},i}^{-2}$, then from Lemma 2.3.2,

$$\begin{aligned}\mathbb{P}(\text{Failure}) \ &\leq \sum_{i=2}^{n}\sum_{t=2}^{T_s}\frac{2\delta}{a_0 n t_{\text{eff},i}^2} \\ &\leq \delta\end{aligned} \tag{A.47}$$

The finiteness of the infinite sum of ${t_{\text{eff},i}}^{-2}$ can be found in Lemma A.7.13.

Thus, in order to keep $\mathbb{P}(\text{Failure}) \leq \delta$, it is sufficient if, at the time of elimination of arm i, we have enough samples to ensure,

$$\begin{aligned}\gamma_i(\boldsymbol{\pi}_T) \ &\leq \frac{\Delta_i^\zeta}{2} \\ \sqrt{\frac{1}{t_{\text{eff},i}}}\left(2\sigma\sqrt{14\log\left(\frac{2a_0 n t_{\text{eff},i}^2}{\delta}\right)} + \rho\epsilon\right) \ &\leq \frac{\Delta_i^\zeta}{2}\end{aligned} \tag{A.48}$$

Rewriting the above equation,

$$\frac{\log(a_i)}{a_i} \ \leq \sqrt{\frac{\delta}{d_1}}\frac{(\Delta_i^\zeta)^2}{d_0} \tag{A.49}$$

where $d_0 = 64 \times 14\sigma^2, d_1 = 2na_0 e^{\frac{\rho^2\epsilon^2}{4\times 14\sigma^2}}$ and $a_i = \sqrt{\frac{d_1}{\delta}} t_{\text{eff},i}$. The following bound on a_i is sufficient to satisfy eq. (A.49),

$$a_i \ \geq 2\sqrt{\frac{d_1}{\delta}} \frac{d_0}{(\Delta_i^\zeta)^2} \log\left(\sqrt{\frac{d_1}{\delta}} \frac{d_0}{(\Delta_i^\zeta)^2}\right)$$

Resubstituting $t_{\text{eff},i}$, we obtain the sufficient number of plays required to eliminate arm i as,

$$t_{\text{eff},i} \ \geq \frac{c_1}{(\Delta_i^\zeta)^2}\left[\log\left(\frac{c_2}{\delta^{\frac{1}{2}}(\Delta_i^\zeta)^2}\right) + c_3\right] \tag{A.50}$$

where $c_1 = 2 \times 64 \times 14\sigma^2$, $c_2 = 64 \times 14\sigma^2\sqrt{2na_0}$ and $c_3 = \frac{\rho^2\epsilon^2}{8\times 14\sigma^2}$.

The further part of the proof depends crucially on the following bound on $t_{\text{eff},i}$ for all $i \in [n]$ from Theorem A.3.1 as follows:

$$t_{\text{eff},i} \geq t_i + \frac{1}{2}\min\left\{\rho\mathfrak{J}(i), T - t_i\right\} \tag{A.51}$$

Hence a sufficiency condition for the GRUB to produce the ζ-best arm with probability $1-\delta$ is given when both the following conditions are satisfied,

$$t_i + \frac{\rho\mathfrak{J}(i)}{2} \geq \frac{1}{(\Delta_i^\zeta)^2}\left[c_1 \log\left(\frac{c_2}{\delta\Delta_i^\zeta}\right) + \frac{\rho\epsilon}{2}\right] \tag{A.52}$$

and,

$$T + t_i \geq T \geq \frac{2}{(\Delta_i^\zeta)^2}\left[c_1 \log\left(\frac{c_2}{\delta\Delta_i^\zeta}\right) + \frac{\rho\epsilon}{2}\right] \tag{A.53}$$

From the Definition A.6.1 we have the set of competitive arms $\mathcal{H}$ and non-competitive arms $\mathcal{N}$ as follows:

$$\mathcal{H} = \left\{j \in [n] \middle| \Delta_i^\zeta \leq 2\sqrt{\frac{2}{\rho\mathfrak{J}(i)}}\left(2\sigma\sqrt{14\log\left(\frac{2a_0 n\rho^2\mathfrak{J}(i)^2}{\delta}\right)} + \rho\epsilon\right)\right\} \tag{A.54}$$

After the first $\max_{i\in\mathcal{N}}\left\{\frac{2}{(\Delta_i^\zeta)^2}\left[c_1\log\frac{c_2}{\delta\Delta_i^\zeta} + \frac{\rho\epsilon}{2}\right]\right\}$ samples, all arms in $\mathcal{N}$ are eliminated. Further, let k_1 be the index of the first arm to be eliminated (in $\mathcal{H}$) and $t^*_{k_1}$ be the number of samples of arm k_1 before getting eliminated then the total number of additional time steps played until the arm k_1 is eliminated is at most $|\mathcal{H}|t^*_{k_1}$. Let k_2 be the index of the next arm in $\mathcal{H}$ to be eliminated. The number of additional plays until the next arm is eliminated is given by $(|\mathcal{H}|-1)[t^*_{k_2} - t^*_{k_1}]$ and so on.

Summing up all the samples required to converge to the optimal arm is given by, (let $t^*_{k_0} = 0$)

$$\sum_{h=1}^{|\mathcal{H}|}(|\mathcal{H}| - h))[t^*_{k_h} - t^*_{k_{h-1}}] = \sum_{h=1}^{|\mathcal{H}|-1} t^*_{k_h} = \sum_{i\in\mathcal{H}/1} t^*_i \tag{A.55}$$

Hence the final sample complexity can be computed as follows:

- Number of plays required for arms in $\mathcal{H}$:

$$\sum_{i\in\mathcal{H}/1} t^*_i \geq \sum_{i\in\mathcal{H}/1} \frac{1}{(\Delta^\zeta_i)^2}\left[c_1 \log \frac{c_2}{\delta\Delta^\zeta_i} + \frac{\rho\epsilon}{2}\right] \tag{A.56}$$

- Number of plays required for all the arms in $\mathcal{N} := [n]/\mathcal{H}$ to be eliminated:

$$T \geq \max_{i\in\mathcal{N}} \left\{\frac{2}{(\Delta^\zeta_i)^2}\left[c_1 \log \frac{c_2}{\delta\Delta^\zeta_i} + \frac{\rho\epsilon}{2}\right]\right\} \tag{A.57}$$

Hence the final sample complexity can be given by,

$$T_{\text{sufficient}} \triangleq \max_{i\in\mathcal{N}} \left\{\frac{2}{(\Delta^\zeta_i)^2}\left[c_1 \log \frac{c_2}{\delta\Delta^\zeta_i} + \frac{\rho\epsilon}{2}\right]\right\} + \sum_{i\in\mathcal{H}/1} \frac{1}{(\Delta^\zeta_i)^2}\left[c_1 \log \frac{c_2}{\delta\Delta^\zeta_i} + \frac{\rho\epsilon}{2}\right] \tag{A.58}$$

□

We extend Lemma A.6.2 to the case when graph G has disconnected clusters.

Theorem A.6.3. *Consider n-armed bandit problem with mean vector $\boldsymbol{\mu} \in \mathbb{R}^n$. Let $\mathcal{G}$ be the set of subgraphs given similarity graph G on the vertex set $[n]$, and further suppose that $\boldsymbol{\mu}$ is ϵ-smooth. Define*

$$\begin{aligned} T_{\textit{sufficient}} \triangleq \min_{D\in\mathcal{G}} \sum_{C\in\mathcal{C}_D} \Bigg[&\sum_{i\in C\cap\mathcal{H}_D} \frac{1}{(\Delta^\zeta_i)^2}\left[c_1 \log \frac{c_2}{\delta\Delta^\zeta_i} + \frac{\rho\epsilon}{2}\right] \\ &+ \max_{i\in C\cap\mathcal{N}_D} \left\{\frac{2}{(\Delta^\zeta_i)^2}\left[c_1 \log \frac{c_2}{\delta\Delta^\zeta_i} + \frac{\rho\epsilon}{2}\right]\right\}\Bigg] \end{aligned} \tag{A.59}$$

where $\Delta^\zeta_i = \max\{\Delta_i, \zeta\}$ for all suboptimal arms, $\mathcal{H}_D$ and $\mathcal{N}_D$ are as in Definition A.6.1, $\mathcal{C}_D$ is the set of connected components of subgraph $D \in \mathcal{G}$ and c_1, c_2 are constants independent of system parameters. Then, with probability at least $1-\delta$, GRUB: (a) terminates in no more than $T_{\textit{sufficient}}$ rounds, and (b) returns a ζ-best arm

Proof. From Lemma A.6.2, the sample complexity for each connected component $C \in \mathcal{C}$ can be given as,

$$T_{\text{sufficient}} = \left[\sum_{i \in C \cap \mathcal{H}} \frac{1}{(\Delta_i^\zeta)^2} \left[c_1 \log \frac{c_2}{\delta \Delta_i^\zeta} + \frac{\rho\epsilon}{2} \right] + \max_{i \in C \cap \mathcal{N}} \left\{ \frac{2}{(\Delta_i^\zeta)^2} \left[c_1 \log \frac{c_2}{\delta \Delta_i^\zeta} + \frac{\rho\epsilon}{2} \right] \right\} \right] \tag{A.60}$$

where, summing it over all the components $C \in \mathcal{C}$, gives us the sample complexity for GRUB while considering graph G.

Any subgraph D of graph G satisfies,

$$\langle \boldsymbol{\mu}, L_G \boldsymbol{\mu} \rangle \leq \epsilon \Rightarrow \langle \boldsymbol{\mu}, L_D \boldsymbol{\mu} \rangle \leq \epsilon \tag{A.61}$$

As seen in Definition A.6.1, the influence factor is instrumental in deciding the competitive and non-competitive sets, which further dictates the sample complexity bounds. Further, notice from Lemma A.7.8 that the influence factor $\mathfrak{I}(i, D)$ is not monotonic when considering subgraph D of graph G. Hence considering a subgraph of G could potentially increase the number of non-competitive arms and provide us with a tighter bound on the performance for GRUB.

Hence $T_{\text{sufficient}}$ can be made tighter by considering the minimum value over the entire set of subgraphs $\mathcal{G}$. □

Note that, as in the case of GRUB, the ζ-GRUB algorithm's performance *automatically* adapts to the best possible subgraph in $\mathcal{G}$.

A.7 Supporting Proofs

This appendix is devoted to providing supporting results for many of the theorems and lemmas in the paper. Let $\{t_i(T)\}_{i=1}^n$ (denoted as $\{t_i\}_{i=1}^n$ for ease of reading) indicate the number of plays of each arm until time T. Let $X \in \mathbb{R}^{n \times n}$ be a matrix, then $\{\lambda_i(X)\}_{i=1}^n$ indicate the eigenvalues of matrix X in an increasing order.

Let $N(\boldsymbol{\pi}_T) = \sum_{t=1}^T \mathbf{e}_{\pi_t} \mathbf{e}_{\pi_t}^T$ be the diagonal counting matrix. Note that $N(\boldsymbol{\pi}_T)$ can be written as $N(\{t_i\}_{i=1}^n)$ since the diagonal counting matrix only depends on the number of plays of each arm, rather than the each sampling sequence $\boldsymbol{\pi}_T$.

We next establish some properties of the influence function $\mathfrak{I}$.

Lemma A.7.1. *Let D be an arbitrary graph with n nodes and let $\{t_i\}_{i=1}^n$ be the number of times all arms are sampled till time T. For each node $j \in [n]$, the following are equivalent:*

$$\begin{aligned}
\frac{1}{\mathfrak{I}(j, D)} &= \max_{\sum_{i \in D_j, i \neq j} t_i = T} \{[K(i, D)]_{jj}\} \qquad (A) \\
&= \max_{k \in D_j, \sum_{i \in D_j, i \neq j} t_i = T} \left\{ [V_j(\{t_i\}_{i \in D_j}, D)^{-1}]_{jj} - [V_j(\{t_i\}_{i \in D_j}, D)^{-1}]_{kk} \right\} \qquad (B) \\
&= \max_{\sum_{i \in D_j, i \neq j} t_i = T} \left\{ [V_j(\{t_i\}_{i \in D_j}, D)^{-1}]_{jj} - \min_{k \in D_j} [V_j(\{t_i\}_{i \in D_j}, D)^{-1}]_{kk} \right\} \qquad (C) \\
&= \max_{\sum_{i \in D_j, i \neq j} t_i = T} \left\{ [V_j(\{t_i\}_{i \in D_j}, D)^{-1}]_{jj} - \frac{1}{T} \right\} \qquad (D)
\end{aligned} \tag{A.62}$$

where $K(i, D)$ be defined as in Definition 2.4.2

Proof. Let $f(\cdot, \cdot)$ denote the following:

$$f(i, D) = \max_{\sum_{i \in D_j, i \neq j} t_i = T} \{[K(i, D)]_{jj}\}$$

We prove the rest by showing equivalence between $(A), (B), (C)$ and (D).

- $(A) \Leftrightarrow (D)$: A simple extension of Lemma A.7.3 to the case of disconnected clustered graph D, $\forall \boldsymbol{\pi}_T \in \mathcal{U}(T, D_j)$ we obtain,

$$V_j(\boldsymbol{\pi}_T, D)^{-1} = \frac{1}{T} \mathbb{1}\mathbb{1}^T + K(\pi_1, D) \tag{A.63}$$

  where $K(\pi_1, D)$ is as defined in Definition 2.4.2. Thus, we have the equivalence by explicitly writing the diagonal element of eq (A.63),

$$[V_j(\boldsymbol{\pi}_T, D)^{-1}]_{jj} - \frac{1}{T} = [K(\pi_1, D)]_{jj} \tag{A.64}$$

  Hence we have the equivalence as,

$$f(i, D) = \max_{\sum_{i \in D_j, i \neq j} t_i = T} \left\{ [V_j(\{t_i\}_{i \in D_j}, D)^{-1}]_{jj} - \frac{1}{T} \right\} \tag{A.65}$$

- $(C) \Leftrightarrow (D)$: Let $\{t_i^*\}_{i \in D_j}$ denote the following:

$$\{t_i^*(j)\}_{i \in D_j} \in \underset{\sum_{i \in D_j, i \neq j} t_i = T}{\arg\max} \left\{ [V_j(\{t_i\}_{i \in D_j}, D)]_{jj}^{-1} - \frac{1}{T} \right\} \tag{A.66}$$

  From Lemma A.7.2, the optimal $\{t_i^*(j)\}_{i \in D_j}$ occurs in $\mathcal{U}(j, T)$, i.e. $\exists \{t_i^*(j)\}_{i \in D_j}$ such that $t_l^*(j) = T$ and $t_k^*(j) = 0 \quad \forall k \neq l$ for some $l \in D_j$. Further by Lemma A.7.4,

$$\min_{k \in D_j} [V_j(\{t_i\}_{i \in D_j}, D)^{-1}]_{kk} = \frac{1}{T} \tag{A.67}$$

  Hence $\{t_i^*(j)\}_{i \in D_j}$ is also a solution for the following problem:

$$\begin{aligned} \{t_i^*(j)\}_{i \in D_j} \in \underset{\sum_{i \in D_j, i \neq j} t_i = T}{\arg\max} & \left\{ [V_j(\{t_i\}_{i \in D_j}, D)]_{jj}^{-1} \right. \\ & \left. - \min_{k \in D_j} [V_j(\{t_i\}_{i \in D_j}, D)^{-1}]_{kk} \right\} \end{aligned} \tag{A.68}$$

  Hence we can conlcude that,

$$\begin{aligned} f(i, D) = \max_{\sum_{i \in D_j, i \neq j} t_i = T} & \left\{ [V_j(\{t_i\}_{i \in D_j}, D)]_{jj}^{-1} \right. \\ & \left. - \min_{k \in D_j} [V_j(\{t_i\}_{i \in D_j}, D)^{-1}]_{kk} \right\} \end{aligned} \tag{A.69}$$

- $(B) \Leftrightarrow (C)$:

  Note that $\max_{k \in D_j, \sum_{i \in D_j, i \neq j} t_i = T} [V_j(\{t_i\}_{i \in D_j}, D)^{-1}]_{jj}]$ does not depend on arm node index $k \in D_j$. Hence, the equivalence follows.

The resistance distance $r(i,j)$ Definition 2.4.1 is independent of δ for all $i, j \in [n]$ (The addition of diagonal elements and subtraction of off diagonal elements removes the dependence on δ Bapat and Gupta (2010)).

Note that $V_T = N_T + \rho L_G$, hence V_T^{-1} gives the psuedo-inverse of the Laplacian matrix for graph G. We show in Lemma A.7.2 that the matrix R (denoting as $R(\delta)$ to explicitly show dependence on δ) linked with V_T^{-1} is independent of number of samples T. Since both matrix R and V_T are psuedo-inverse of the Laplacian L_G. Thus we can conclude the following :

$$\lim_{\delta \to 0} [R(\delta)]_{ij} - \frac{1}{\delta} = \lim_{T \to 0} [V(\{t_i\}_{i=1}^n, G)^{-1}]_{ij} - \frac{1}{T} \tag{A.70}$$

where $T \to 0$ implies $t_i \to 0 \quad \forall i \in [n]$. Further,

$$\begin{aligned} &\lim_{\delta \to 0} R(\delta)_{ii} + R(\delta)_{jj} - R(\delta)_{ij} - R(\delta)_{ji} \\ &= \lim_{T \to 0} [V(\{t_i\}_{i=1}^n, G)^{-1}]_{ii} + [V(\{t_i\}_{i=1}^n, G)^{-1}]_{jj} \\ &- [V(\{t_i\}_{i=1}^n, G)^{-1}]_{ij} - [V(\{t_i\}_{i=1}^n, G)^{-1}]_{ji} \end{aligned} \tag{A.71}$$

where $T \to 0$ implies $t_i \to 0 \quad \forall i \in [n]$.

Since the equation (A.71) holds for $t_i \to 0$ for all $i \in [n]$, computing the value of limit for one trajectory should suffice for finding the value of the limit. Thereby, we provide an alternate equation for obtaining the resistance distance $r(i,j)$ by

$$r(i,j) = [K(\pi_1 = i, D)]_{jj} \tag{A.72}$$

Note that $[K(\pi_1 = i, D)]_{ii} = [K(\pi_1 = i, D)]_{ij} = [K(\pi_1 = i, D)]_{ji} = 0$ from Lemma A.7.3). Thus we can say from Definition 2.4.2,

$$f(i, D) = \frac{1}{\mathfrak{I}(j, D)}$$

Hence proved.

□

Lemma A.7.2. *Let D be a given graph with n nodes. For every node $j \in D$, let $\{t_i^*(j)\}_{i \in D_j}$ denote the following:*

$$\{t_i^*(j)\}_{i \in D_j} \in \underset{\sum_{i \in D_j, i \neq j} t_i = T}{\arg\max} \left\{ [V_j(\{t_i\}_{i \in D_j}, D)]_{jj}^{-1} - \frac{1}{T} \right\} \tag{A.73}$$

Then $\exists \{t_i^*(j)\}_{i \in D_j}$, $l \in D_j$ such that $t_l^*(j) = T$ and $t_k^*(j) = 0 \quad \forall k \neq l$.

Proof. To simplify our proof, let graph D be connected. The proof for the case of disconnected components is an extension of the connected graph case, by analysing each individual connected component together.

If graph D is connected then $D_i = D$. For the rest of the proof we sometimes denote $V(\boldsymbol{\pi}_T, D)$ as $V(\{t_i\}_{i=1}^n, D)$ to make it more context relevant.

Let $g : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ be a partial function of $V(\boldsymbol{\pi}_T, D)$ as follows:

$$g(\{t_i\}_{i=1}^n) = V(\{t_i\}_{i=1}^n, D) \tag{A.74}$$

For all $i \in [n]$, let $t_i = \alpha_i T$ such that $\sum_{i=1}^n \alpha_i = 1$. Then we can say that,

$$\begin{aligned} g(\{t_i\}_{i=1}^n) &= g(\{\alpha_i T\}_{i=1}^n) \\ &= \sum_{i=1}^n \alpha_i g(\{0, 0, \dots t_i = T, \dots 0\}) \end{aligned} \tag{A.75}$$

Using convexity of matrix invertibility Nordström (2011) $V(\boldsymbol{\pi}_T, G)^{-1}$ satisfies,

$$g(\{t_i\}_{i=1}^n)^{-1} \preceq \sum_{i=1}^n \alpha_i g(\{0, 0, \dots t_i = T, \dots 0\})^{-1} \tag{A.76}$$

Hence $g(\cdot)^{-1}$ is a convex function. Since we have the restriction as $\sum_{i=1, i\neq j}^n t_i = T$. We can say that,

$$\begin{aligned} &\underset{\sum_{i\in D_j, i\neq j} t_i = T}{\arg\max} \left\{ [V(\{t_i\}_{i=1}^n, D)]_{jj}^{-1} - \frac{1}{\sum_{i=1}^n t_i} \right\} \\ &= \underset{\sum_{i\in D_j, i\neq j} t_i = T}{\arg\max} [V(\{t_i\}_{i=1}^n, D)]_{jj}^{-1} \\ &= \underset{\sum_{i\in D_j, i\neq j} t_i = T}{\arg\max} \langle \mathbf{e}_j, [V(\{t_i\}_{i=1}^n, D)]^{-1} \mathbf{e}_j \rangle \\ &= \underset{\sum_{i\in D_j, i\neq j} t_i = T}{\arg\max} \langle \mathbf{e}_j, g(\{t_i\}_{i=1}^n)^{-1} \mathbf{e}_j \rangle \end{aligned} \tag{A.77}$$

Since $g(\cdot)^{-1}$ is convex, for a convex function the maximization over a simplex happens at one of the vertices. Hence the max happens when $t_i = T$ and $t_k = 0 \quad \forall k \neq i$.

Hence proved. □

Lemma A.7.3. *Let G be a given connected graph of n nodes and t_i be the number of samples of each arm i. Then $\forall \boldsymbol{\pi}_T \in \mathcal{U}(T)$,*

$$V(\boldsymbol{\pi}_T, G)^{-1} = \frac{1}{T} \mathbb{1}\mathbb{1}^T + K(\pi_1, G) \tag{A.78}$$

where, $\mathbb{1} \in \mathbb{R}^n$ is a vector or all ones and $K(\pi_1, G) \in \mathbb{R}^{n\times n}$ is the matrix defined in Definition 2.4.2.

Proof. Let I be an identity matrix of dimension $n \times n$. We prove the result by showing that, $\forall \boldsymbol{\pi}_T \in \mathcal{U}(T)$, $V(\boldsymbol{\pi}_T, G)^{-1} V(\boldsymbol{\pi}_T, G) = I$,

$$\begin{aligned} & V(\boldsymbol{\pi}_T, G)^{-1} V(\boldsymbol{\pi}_T, G) \\ &= \left(\frac{1}{T} \mathbb{1}\mathbb{1}^T + K(\pi_1, G) \right) \left(\sum_{t=1}^{T} \mathbf{e}_{\pi_t} \mathbf{e}_{\pi_t}^T + \rho L_G \right) \\ &= \left(\frac{1}{T} \mathbb{1}\mathbb{1}^T + K(\pi_1, G) \right) \left(T \mathbf{e}_{\pi_1} \mathbf{e}_{\pi_1}^T + \rho L_G \right) \\ &= \mathbb{1}\mathbf{e}_{\pi_1}^T + T K(\pi_1, G) \mathbf{e}_{\pi_1} \mathbf{e}_{\pi_1}^T + \rho K(\pi_1, G) L_G \end{aligned} \tag{A.79}$$

From Definition 2.4.2, $K(\pi_1, G)\mathbf{e}_{\pi_1}\mathbf{e}_{\pi_1}^T = 0$ and $\mathbb{1}\mathbf{e}_{\pi_1}^T + \rho K(\pi_1, G) L_G = I$ implying that $V(\boldsymbol{\pi}_T, G)^{-1} V(\boldsymbol{\pi}_T, G) = I$.

Hence proved. □

Lemma A.7.4. *Let G be any connected graph and $\boldsymbol{\pi}_T \in \mathcal{U}(T, G)$. Then,*

$$\min_{j \in [n]} \{ [V(\boldsymbol{\pi}_T, G)^{-1}]_{jj} \} = \frac{1}{T} \tag{A.80}$$

Proof. From Definition 2.4.2, $K(\pi_1, G)$ satisfies

$$K(\pi_1, G) L_G = \frac{1}{\rho} \left(I - \mathbb{1}\mathbf{e}_{\pi_1}^T \right)$$

Observe that $\mathbb{1}\mathbf{e}_i^T$ is a rank 1 matrix with eigenvalue 1 and eigenvector $\mathbf{e}_i$ and Identity matrix I is of rank n with all eigenvalues 1 and eigenvectors $\{\mathbf{e}_i\}_{i=1}^n$. Hence $\left(I - \mathbb{1}\mathbf{e}_{\pi_1}^T \right)$ is a rank $n-1$ matrix with rest nonzero eigenvalues as 1. Since the graph G is connected, $\lambda_1(L_G) = 0$ and $\lambda_2(L_G) > 0$. The eigenvector corresponding to $\lambda_1(L_G)$ is $\mathbb{1}$, the all 1 vector.

Given $\rho > 0$, we can conclude,

$$K(\pi_1, G) L_G \succeq 0 \quad \text{s.t. } \operatorname{rank}(K(\pi_1, G) L_G) = n - 1 \tag{A.81}$$

Hence, in order to satisfy eq. (A.81), $K(\pi_1, G) \succeq 0$ and $\operatorname{rank}(K(\pi_1, G)) \geq n - 1$. By lower bounds on Rayleigh quotient we can conclude,

$$\langle \mathbf{e}_j, K(\pi_1, G) \mathbf{e}_j \rangle = [K(\pi_1, G)]_{jj} \geq 0 \quad \forall j \in [n] \tag{A.82}$$

From Lemma A.7.3, $[K(\pi_1, G)]_{jj} = [V(\boldsymbol{\pi}_T, G)^{-1}]_{jj} - \frac{1}{T}$ implying that $[V(\boldsymbol{\pi}_T, G)^{-1}]_{jj} \geq \frac{1}{T}$. From Definition 2.4.2 it can be seen that $[K(\pi_1, G)]_{\pi_1 \pi_1} = 0$ and hence $[V(\boldsymbol{\pi}_T, G)^{-1}]_{\pi_1 \pi_1} = \frac{1}{T}$ which concludes the proof. □

Lemma A.7.5. *Given a connected graph G, the following bound holds for all the diagonal entries of $[V(\boldsymbol{\pi}_T, G)^{-1}]_{ii}$ for $i \in [n]$:*

$$[V(\boldsymbol{\pi}_T, G)^{-1}]_{ii} \leq \mathbb{1}(t_i = 0) \left(\frac{1}{\rho \mathfrak{I}(i, \mathcal{G})} + \frac{1}{T} \right) + \mathbb{1}(t_i > 0) \max \left\{ \frac{1}{t_i + \frac{\rho \mathfrak{I}(i,G)}{2}}, \frac{1}{t_i + \frac{T}{2}} \right\} \tag{A.83}$$

Proof. From Definition 2.4.2 of $\mathfrak{I}(\cdot, \mathcal{G})$ and Lemma A.7.1, Breaking the lemma statement into cases:

- **Unsampled Arms :** From Lemma A.7.1

$$\frac{1}{\mathfrak{I}(j,G)} = \max_{\sum_{i\in G_j, i\neq j} t_i = T} \left\{ [V_j(\{t_i\}_{i\in G_j}, G)^{-1}]_{jj} - \frac{1}{T} \right\} \quad \forall j \in [n] \tag{A.84}$$

  Thus for any unsampled arm j,

$$[V(\boldsymbol{\pi}_T, G)]^{-1}_{jj} \leq \left(\frac{1}{\mathfrak{I}(j,G)} + \frac{1}{T} \right) \tag{A.85}$$

- **Sampled Arms :** Since the matrix $V(\boldsymbol{\pi}_T, G)$ depends only on the final sampling distribution $\{t_i\}_{i=1}^n$ rather than the sampling path $\boldsymbol{\pi}_T$. Consider a sampling path such that $\pi_t \neq j$ for $t \leq T - t_j$ and $\pi_t = j$ for $T - t_j \leq t \leq T$.

  Assuming such a sampling path $\boldsymbol{\pi}_T$, after $\boldsymbol{\pi}_{T-t_j}$ samples,

$$[V(\boldsymbol{\pi}_{T-t_j}, G)^{-1}]_{jj} \leq \frac{1}{T} + \frac{1}{\mathfrak{I}(j,G)} \tag{A.86}$$

  Then by the Sherman-Morrison rank 1 update identity (Hager, 1989),

$$\begin{aligned}
\frac{1}{[V(\boldsymbol{\pi}_T, G)^{-1}]_{jj}} &= \frac{1}{[V(\boldsymbol{\pi}_{T-t_j}, G)^{-1}]_{jj}} + t_j \\
[V(\boldsymbol{\pi}_T, G)^{-1}]_{jj} &= \frac{1}{t_j + \frac{1}{[V(\boldsymbol{\pi}_{T-t_j}, G)^{-1}]_{jj}}} \\
&\leq \frac{1}{t_j + \frac{1}{\left(\frac{1}{\mathfrak{I}(j,G)} + \frac{1}{T-t_j}\right)}}
\end{aligned}$$

  Hence we have the bound on $[V(\boldsymbol{\pi}_T, G)^{-1}]_{jj}$ as follows:

$$[V(\boldsymbol{\pi}_T, G)^{-1}]_{jj} \leq \max \left\{ \frac{1}{t_j + \frac{\mathfrak{I}(j,G)}{2}}, \frac{1}{t_j + \frac{T-t_j}{2}} \right\} \tag{A.87}$$

Hence proved. □

Lemma A.7.6. *Let D be a graph with n nodes and k disconnected components. If each of the connected components $\{\mathcal{C}_i(D)\}_{i=1}^k$ is a complete graph then $\forall\ j \in [n]$,*

$$\mathfrak{I}(j, D) = \frac{|\mathcal{C}(j, D)|}{2} \tag{A.88}$$

Proof. Let D be a complete graph (k = 1), $\boldsymbol{\pi}_T \in \mathcal{U}(T)$ and $\rho = 1$. Then,

$$V(\boldsymbol{\pi}_T, G)^{-1} = \frac{1}{T}\mathbb{1}\mathbb{1}^T + K \tag{A.89}$$

where $\mathbb{1} \in \mathbb{R}^n$ is a vector or all ones and $K \in \mathbb{R}^{n\times n}$ is a matrix given by,

$$K_{\pi_1\pi_1} = 0, \quad K_{jj} = \frac{2}{n} \;\; \forall j \in [n]/\{\pi_1\}$$
$$K_{k\pi_1} = 0, \quad K_{\pi_1 j} = 0, \quad K_{jk} = \frac{1}{n} \;\; \forall j,k \in [n]/\{\pi_1\},\; j \neq k$$

The form of $V(\boldsymbol{\pi}_T, G)^{-1}$ in eq.(A.89) can be verified by $V(\boldsymbol{\pi}_T, G)^{-1}V(\boldsymbol{\pi}_T, G) = I$.

The final statement of the lemma can be obtained by considering this analysis to just the nodes within a connected component of a diconnected graph G and Lemma A.7.1. □

Lemma A.7.7. *Let D be a graph with n nodes and k disconnected components. If each of the connected components $\{\mathcal{C}_i(D)\}_{i=1}^k$ is a line graph then $\forall\; j \in [n]$,*

$$\mathfrak{I}(j, D) > \frac{1}{|\mathcal{C}(j, D)|} \tag{A.90}$$

Proof. Let D be a complete graph (k = 1), $\boldsymbol{\pi}_T \in \mathcal{U}(T)$ and $\rho = 1$. Then,

$$V(\boldsymbol{\pi}_T, G)^{-1} = \frac{1}{T}\mathbb{1}\mathbb{1}^T + K \tag{A.91}$$

where $\mathbb{1} \in \mathbb{R}^n$ is a vector or all ones and $K \in \mathbb{R}^{n\times n}$ is a matrix given by,

$$K_{\pi_1\pi_1} = 0, \quad K_{jj} = d(\pi_1, j) \;\; \forall j \in [n]/\{\pi_1\},$$
$$K_{k\pi_1} = 0, \quad K_{\pi_1 j} = 0,$$
$$K_{jk} = \min\{d(\pi_1, j), d(\pi_1, k)\} \;\; \forall j,k \in [n]/\{\pi_1\},\; j \neq k$$

The form of $V(\boldsymbol{\pi}_T, G)^{-1}$ in eq.(A.91) can be verified by $V(\boldsymbol{\pi}_T, G)^{-1}V(\boldsymbol{\pi}_T, G) = I$.

The final statement of the lemma can be obtained by considering this analysis to just the nodes within a connected component of a diconnected graph G and Lemma A.7.1. □

Lemma A.7.8. *Let $A = ([n], E)$ be any graph and let $e \in E$ be an edge of graph A. Let $B = ([n], E - \{e\})$ be a subgraph of A with one edge removed. Then the following holds for all non-isolated nodes i in B:*

- *If $|\mathcal{C}(A)| = |\mathcal{C}(B)|$,*

$$\mathfrak{I}(i, A) \geq \mathfrak{I}(i, B)$$

- *If $|\mathcal{C}(A)| < |\mathcal{C}(B)|$,*

$$\mathfrak{I}(i, A) \leq \mathfrak{I}(i, B)$$

Proof. From Lemma A.7.1, for any graph D, $\mathfrak{I}(\cdot,\cdot)$ satisfies,

$$\begin{aligned}\frac{1}{\mathfrak{I}(j,D)} &= \max_{k\in D_j, \sum_{i\in D_j} t_i = T} \big\{ [V_j(\{t_i\}_{i\in D_j}, D)^{-1}]_{jj} \\ &\quad - [V_j(\{t_i\}_{i\in D_j}, D)^{-1}]_{kk} \big\} \quad \forall j \in [n] \end{aligned} \tag{A.92}$$

Case I : $|\mathcal{C}(A)| = |\mathcal{C}(B)|$

The edge set of B is smaller than edge set of A. Hence, from Lemma $\mathfrak{I}(i,A) \geq \mathfrak{I}(i,B)$

Case II : $\mathcal{C}(A) < \mathcal{C}(B)$ In this case, $|B_i| \leq |A_i|$. Hence the max is over a smaller set of options, we can conclude that $\mathfrak{I}(i,A) \leq \mathfrak{I}(i,B)$. Hence proved. □

Given a graph D, we define a class of sampling policies $\mathcal{U}(T,D)$ as follows,

Definition A.7.9. *Let $\mathcal{U}(T,D)$ denote the set of sampling policies,*

$$\mathcal{U}(T,D) = \{\boldsymbol{\pi}_T |\ \exists l \in D \ \ s.t.\ \pi_t = l \ \ \forall t \leq T\}$$

Lemma A.7.10. *Let G be the given graph and sampling policy $\boldsymbol{\pi}_T$ has been played for T time steps, then V_T satisfies the following structure,*

$$V(\boldsymbol{\pi}_T, D) = diag([V_1, V_2, \ldots, V_{k(G)}]) \tag{A.93}$$

where V_i depends on the connected component $C_i \in \mathcal{C}_D$ of the graph and the number of samples of the arms within the connected component $\{t_j\}_{j\in C_i}$.

Proof. Rewriting the definition of $V(\boldsymbol{\pi}_T, D)$,

$$\begin{aligned} V(\boldsymbol{\pi}_T, D) &\triangleq \sum_{t=1}^{T} \mathbf{e}_{\pi_t} \mathbf{e}_{\pi_t}^\top + \rho L_D \\ &= N(\{t_i\}_{i=1}^n) + L_D \end{aligned} \tag{A.94}$$

Both component matrices $N(\{t_i\}_{i=1}^n)$ (diagonal matrix) and L_D (Laplacian matrix of a graph) adhere to a block diagonal structure and hence $V(\boldsymbol{\pi}_T, D)$ matrix also adheres to a block diagonal structure analogous to L_D. The block diagonal structure in L_D is dictated by connected components of graph D. □

The following lemma establishes the invertibility of $V(\boldsymbol{\pi}_T, G)$ for a connected graph and $T > 1$:

Lemma A.7.11. *For a connected graph G, $V(\boldsymbol{\pi}_1, G)$ is invertible, but $V(\boldsymbol{\pi}_0, G)$ is not invertible.*

Proof. Since the graph G is connected, $\lambda_1(L_G) = 0$ and $\lambda_2(L_G) > 0$. The eigenvector corresponding to $\lambda_1(L_G)$ is $\mathbb{1}$, the all 1 vector. At time $T = 0$, $V(\boldsymbol{\pi}_T, G) = L_G$ and hence $V(\boldsymbol{\pi}_T, G)$ is positive semi-definite matrix with one zero eigenvalues.

Let arm i be pulled at $T = 1$, i.e. $\pi_1 = i$, then the corresponding counting matrix is a positive semi definite matrix of rank one with the eigen value $\lambda_n(N) = 1$ for the eigenvector e_i.

Observe that $\mathbf{e}_i^T \mathbb{1} > 0$. Also, N_T and L_G are positive semi-definite matrices with ranks 1 and $n-1$ respectively. The subspace without information (corresponding to the direction of zero eigenvalue) for matrix L_G is now provided by $N(\boldsymbol{\pi}_1)$ and hence $\lambda_{\min}(V(\boldsymbol{\pi}_1, G)) > 0$ making it invertible. □

Lemma A.7.12. *Let $G = ([n], E_G, A), H = ([n], E_H, A)$ are two graphs with n nodes such that $E_G \supseteq E_H$. Then, assuming invertibility of $[V(G,T)^{-1}]$ and $[V(H,T)^{-1}]$,*

$$[t_{\mathit{eff},i}]_G \geq [t_{\mathit{eff},i}]_H \quad \forall i \in [n], T > k(G) \tag{A.95}$$

where $\forall i \in [n]$, $[t_{\mathit{eff},i}]_G, [t_{\mathit{eff},i}]_H$ indicates the effective samples with graph G and H respectively.

Proof. Given graphs $G = ([n], E_G), H = ([n], E_H)$ satisfy $E_G \supseteq E_H$.

The quadratic form of Laplacian for the graph G, H is given by,

$$\mathbf{x}L_G\mathbf{x} = \sum_{(i,j)\in E_G} (x_i - x_j)^2$$
$$\mathbf{x}L_H\mathbf{x} = \sum_{(i,j)\in E_H} (x_i - x_j)^2$$

Since $E_G \supseteq E_H$,

$$\mathbf{x}L_G\mathbf{x} \geq \mathbf{x}L_H\mathbf{x} \quad \forall\ \mathbf{x} \in \mathbb{R}^n$$
$$\Rightarrow L_G \succeq L_H$$

Further, provided a sampling policy $\boldsymbol{\pi}_T$, we can say that,

$$V(\boldsymbol{\pi}_T, G) \succeq V(\boldsymbol{\pi}_T, H)$$

For the number of samples T sufficient to ensure invertibility of $V(\boldsymbol{\pi}_T, H)$, we have

$$V(\boldsymbol{\pi}_T, G)^{-1} \preceq V(\boldsymbol{\pi}_T, H)^{-1}$$
$$\mathbf{x}^T V(\boldsymbol{\pi}_T, G)^{-1}\mathbf{x} \leq \mathbf{x}^T V(\boldsymbol{\pi}_T H)^{-1}\mathbf{x} \quad \forall \mathbf{x} \in \mathbb{R}^n$$
$$[V(\boldsymbol{\pi}_T, G)^{-1}]_{ii} \leq [V(\boldsymbol{\pi}_T, H)^{-1}]_{ii} \quad (\text{taking } \mathbf{x} = \mathbf{e}_i)$$
$$\frac{1}{[V(\boldsymbol{\pi}_T, G)^{-1}]_{ii}} \geq \frac{1}{[V(\boldsymbol{\pi}_T, H)^{-1}]_{ii}}$$

Hence from the definition of effective samples 2.3.1, it is clear that for any $i \in [n]$,

$$[t_{\text{eff},i}]_G \geq [t_{\text{eff},i}]_H \tag{A.96}$$

Hence proved. □

Lemma A.7.13. *Let effective samples $t_{\mathit{eff},i}$ be as is defined in Definition 2.3.1 and let $\boldsymbol{\pi}_T$ denote a cyclic sampling policy for $T > k(G)$, then the infinite sum $\sum_{T=k(G)+1}^{\infty} t_{\mathit{eff},i}^{-2}$ is bounded. In fact,*

$$\sum_{T=k(G)+1}^{\infty} t_{\mathit{eff},i}^{-2} < n\left(\frac{2(n-1)}{\rho}\right)^2 + \frac{n\pi^2}{6} \tag{A.97}$$

Proof. We first prove the lemma statement for connected graph G and then go towards a more general graph G. From Lemma A.3.1,

$$t_{\text{eff},i} \geq t_i + \min\{\rho\mathfrak{I}(i,G), T - t_i\}$$

if $T - t_i \leq \rho\mathfrak{I}(i,G)$, then $t_{\text{eff},i} \geq \frac{T+t_i}{2} \geq \frac{T}{2}$. For the reverse case of $T - t_i \geq \rho\mathfrak{I}(i,G)$, $t_{\text{eff},i} \geq t_i + \frac{\rho\mathfrak{I}(i,G)}{2} \geq t_i + \frac{\rho}{2(n-1)}$ (since $\mathfrak{I}(i,G) \geq \frac{1}{n-1}$ by Lemma A.7.5).

Since $\boldsymbol{\pi}_T$ is a cyclic sampling policy, hence t_i increases by 1 at-least once every n samples. Thus, we can upperbound the infinite sum as,

$$\begin{aligned}
\sum_{T=1}^{\infty} \frac{1}{t_{\text{eff},i}^2} &\leq \sum_{T=1}^{\infty} \frac{1}{\left(t_i + \frac{\rho}{2(n-1)}\right)^2} \\
&\leq n\left(\frac{2(n-1)}{\rho}\right)^2 + n\sum_{t_i=1}^{\infty} \frac{1}{t_i^2} \\
&< n\left(\frac{2(n-1)}{\rho}\right)^2 + \frac{n\pi^2}{6} \qquad \text{(A.98)}
\end{aligned}$$

Hence proved. □

APPENDIX B

QUADRATIC FEASIBILITY

B.1 Appendix A : Injectivity

B.1.1 Proof of Lemma 3.5.2

The following result from Wang and Xu (2017) is quite crucial for the further proof, restating,

Theorem B.1.2 (Theorem 2.1, Wang and Xu (2017)). *Let $\mathcal{A} = \{A_i\}_{i=1}^m \subset \mathbf{H}_n(\mathbb{C})$. The following statements are equivalent:*

1. *For a given $\mathcal{A} = \{A_i\}_{i=1}^m$, the mapping $\mathcal{M}_\mathcal{A}$ has phase retrieval property.*
2. *There exists no nonzero vector $\mathbf{v}, \mathbf{u} \in \mathbb{C}^n$ with $\mathbf{u} \neq ic\mathbf{v}$, $c \in \mathbb{R}$, such that $Re(\langle A_j\mathbf{u}, \mathbf{v}\rangle) = 0$ for all $1 \leq j \leq m$.*

Lemma B.1.3. *The following statements are equivalent:*

1. *The nonlinear map $\mathcal{M}_\mathcal{A} : \widehat{\mathbb{C}}^n \to \mathbb{C}^m$ is injective.*
2. *There exist constants $\alpha, \beta > 0$ such that $\forall \mathbf{u}, \mathbf{v} \in \mathbb{C}^n$,*

$$\beta\|[[\mathbf{u}, \mathbf{v}]]\|_F^2 \geq \sum_{i=1}^m |\langle A_i, [[\mathbf{u}, \mathbf{v}]]\rangle|^2 \geq \alpha\|[[\mathbf{u}, \mathbf{v}]]\|_F^2.$$

Proof. (1→2)
For the mapping $\mathcal{M}_\mathcal{A}$ to be injective, the following should holds,

$$\mathcal{M}_\mathcal{A}(\mathbf{x}) = \mathcal{M}_\mathcal{A}(\mathbf{y}) \quad \text{iff} \quad \mathbf{x} \sim \mathbf{y}$$

Hence for $\mathbf{x} \nsim \mathbf{y}$, $\mathcal{M}_\mathcal{A}(\mathbf{x}) \neq \mathcal{M}_\mathcal{A}(\mathbf{y})$. It can be verified that for any $\phi \in [0, 2\pi]$, $\mathbf{u} = \mathbf{x} - e^{i\phi}\mathbf{y}, \mathbf{v} = \mathbf{x} + e^{i\phi}\mathbf{y}$ satisfies the following transformation,

$$(\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*) = (\mathbf{u}\mathbf{v}^* + \mathbf{v}\mathbf{u}^*) = [[\mathbf{u}, \mathbf{v}]] \tag{B.1}$$

Thus,

$$\|\mathcal{M}_\mathcal{A}(\mathbf{x}) - \mathcal{M}_\mathcal{A}(\mathbf{y})\|_2^2 = \sum_{i=1}^m |\langle A_i, [[\mathbf{u}, \mathbf{v}]]\rangle|^2$$

We define the lower bound α and upper bound β as below,

$$\alpha := \min_{T \in S^{1,1}, \|T\|_F = 1} \sum_{i=1}^m |\langle A_i, T\rangle|^2,$$
$$\beta := \max_{T \in S^{1,1}, \|T\|_F = 1} \sum_{i=1}^m |\langle A_i, T\rangle|^2$$

The set $T \in S^{1,1}, \|T\|_F = 1$ is compact, hence the constants α, β exists.

From Theorem B.1.2 statement, it is clear that $\langle A_i, [[\mathbf{u}, \mathbf{v}]]\rangle = \langle A_i, \mathbf{u}\mathbf{v}^*\rangle + \langle A_i, \mathbf{v}\mathbf{u}^*\rangle = Re(\langle A_i\mathbf{v}, \mathbf{u}\rangle) = 0$ cannot be satisfied for all $i \in [1, m]$ if $\mathbf{x} \nsim \mathbf{y}$ satisfying equation (B.1).

(1$\leftarrow$2)

Instead of proving 1$\leftarrow$2 , we argue the negation holds, i.e. 1$\nrightarrow$2.

Suppose the mapping $\mathcal{M}_{\mathcal{A}}$ is not injective.

Then $\exists \mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ such that,

$$\mathbf{x} \nsim \mathbf{y}, \quad \mathcal{M}_{\mathcal{A}}(\mathbf{y}) = \mathcal{M}_{\mathcal{A}}(\mathbf{x})$$

Thus $\|\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\|_F \neq 0$, but $\|\mathcal{M}_{\mathcal{A}}(\mathbf{y}) - \mathcal{M}_{\mathcal{A}}(\mathbf{x})\|_2 = 0$. Thus $\alpha = 0$ and hence the negation follows. □

Using the result from Lemma B.1.3, we are now equipped to prove Lemma 3.5.2.

Lemma B.1.4. *The mapping $\mathcal{M}_{\mathcal{A}}$ is injective iff it is (α, β)-stable for some constants $0 < \alpha \leq \beta$.*

Proof. In order to prove the theorem statement, we examine the properties of the ratio,

$$V(\mathbf{x}, \mathbf{y}) = \frac{\sum_{i=1}^{m} |\langle A_i, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2}{\|\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\|_F^2}$$

The (α, β)-stability of the mapping $\mathcal{M}_{\mathcal{A}}$ directly follows from Lemma B.1.3 and the existence of $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ satisfying equation (B.1). □

Theorem B.1.5. *Consider $\mathcal{A} = \{A_i\}_{i=1}^m$ and the mapping $\mathcal{M}_{\mathcal{A}}$ be defined as in* (3.1)*. Viewing $\{A_i\mathbf{x}\}_{i=1}^m$ as vectors in $\mathbb{R}^{2n}$, we can say the following:*

1. *$\mathcal{M}_{\mathcal{A}}$ is injective.*
2. *$\forall I \subset [m]$, then either dim(span($\{A_i\mathbf{x}\}_{i\in I}$))$\geq 2n-1$ or dim(span($\{A_i\mathbf{x}\}_{i\in I^C}$))$\geq 2n-1$ for all vectors $\mathbf{x} \in \mathbb{C}^n \backslash 0$*
3. *$\forall I \subset [m]$, then either span($\{A_i\mathbf{x}\}_{i\in I}$) = span($\{i\mathbf{x}\}$)$^\perp$ or span($\{A_i\mathbf{x}\}_{i\in I^C}$) = span($\{i\mathbf{x}\}$)$^\perp$ for all vectors $\mathbf{x} \in \mathbb{C}^n \backslash 0$.*

Proof. 2 $\leftarrow$ 3: immediately follows

2 $\rightarrow$ 3: Given span$(\{A_i\mathbf{x}\}_{i\in I})_{\mathbb{R}} \geq 2n-1$, note that

$$\langle A_i\mathbf{x}, i\mathbf{x}\rangle_{\mathbb{R}} = (i\langle A_i\mathbf{x}, \mathbf{x}\rangle)_{\mathbb{R}} = 0$$

since $\text{Im}(\langle A_i\mathbf{x}, \mathbf{x}\rangle) = 0$. Thus we can conclude that span$\{i\mathbf{x}\}_{\mathbb{R}} \subset$span$(\{A_i\mathbf{x}\}_{i\in I})^\perp$. Further, given that span$(\{A_i\mathbf{x}\}_{i\in I})^\perp \leq 1$, the relation is proved.

1 $\rightarrow$ 3: We need to show that the spanned space is span$(\{i\mathbf{x}\})^\perp$ i.e. orthogonal space to $i\mathbf{x}$. Consider the following:

$$\langle A(\mathbf{x}+\mathbf{z}), \mathbf{x}+\mathbf{z}\rangle - \langle A(\mathbf{x}-\mathbf{z}), \mathbf{x}-\mathbf{z}\rangle = 4Re\langle A\mathbf{x}, \mathbf{z}\rangle \tag{B.2}$$

Given that the mapping $\mathcal{M}_{\mathcal{A}}$ is injective, equation (B.2) can be 0 if and only if $\mathbf{x}+\mathbf{z} = c(\mathbf{x}-\mathbf{z})$ such that $c \in \mathbf{C}, |c| = 1$. This gives us,

$$\mathbf{x}+\mathbf{z} = c(\mathbf{x}-\mathbf{z}) \quad \Rightarrow \quad \mathbf{z} = -\frac{1-c}{1+c}\mathbf{x} = -\frac{2\text{Im}c}{|1+c|^2}i\mathbf{x} \tag{B.3}$$

This means $(\text{span}_{\mathbb{R}}\{i\mathbf{x}\}^{\perp})^{\perp} \subset \text{span}_{\mathbb{R}}\{i\mathbf{x}\}$. To prove $\text{span}_{\mathbb{R}}\{i\mathbf{x}\} \subset (\text{span}_{\mathbb{R}}\{i\mathbf{x}\}^{\perp})^{\perp}$ take $\mathbf{z} = i\alpha\mathbf{x}$ and $c = \frac{1+\alpha i}{1-\alpha i}$. It can be seen that $|c| = 1$. Then,

$$\begin{aligned} \mathbf{x}+\mathbf{z} &= (1+i\alpha)\mathbf{x} \\ &= \frac{1+i\alpha}{1-i\alpha}(\mathbf{x}-\mathbf{z}) \\ &= c(\mathbf{x}-\mathbf{z}) \end{aligned} \tag{B.4}$$

Thus $\text{span}_{\mathbb{R}}\{i\mathbf{x}\} \subset (\text{span}_{\mathbb{R}}\{i\mathbf{x}\}^{\perp})^{\perp}$.
$3 \rightarrow 1$: Suppose $\mathcal{A}(\mathbf{x}) = \mathcal{A}(\mathbf{y})$ and $\mathbf{x} = c\mathbf{y}, c \in \mathbb{C}, |c| = 1$ then we are done.
If $\mathcal{A}(\mathbf{x}) = \mathcal{A}(\mathbf{y})$ and $\mathbf{x} \neq c\mathbf{y}, c \in \mathbb{C}, |c| = 1$ then note that,

$$\begin{aligned} &\langle A_i(\mathbf{x}-\mathbf{y}), (\mathbf{x}+\mathbf{y})\rangle_{\mathbb{R}} \\ &= \langle A_i\mathbf{x}, \mathbf{x}\rangle_{\mathbb{R}} + \langle A_i\mathbf{x}, \mathbf{y}\rangle_{\mathbb{R}} - \langle A_i\mathbf{y}, \mathbf{x}\rangle_{\mathbb{R}} - \langle A_i\mathbf{y}, \mathbf{y}\rangle_{\mathbb{R}} \\ &= Re\left(\langle A_i\mathbf{x}, \mathbf{y}\rangle - \langle A_i\mathbf{y}, \mathbf{x}\rangle\right) \\ &= Re\left(\langle A_i\mathbf{x}, \mathbf{y}\rangle - \overline{\langle A_i\mathbf{x}, \mathbf{y}\rangle}\right) \\ &= 0 \end{aligned} \tag{B.5}$$

Thus from 3), $\mathbf{x}+\mathbf{y} \in S(\mathbf{x}-\mathbf{y})^{\perp} = \text{span}_{\mathbb{R}}\{i(\mathbf{x}-\mathbf{y})\}$. Further, there exists $\alpha \in \mathbb{C}, |\alpha| = 1$ such that $\mathbf{x}+\mathbf{y} = i\alpha(\mathbf{x}-\mathbf{y})$. After rearranging, we have,

$$\mathbf{x} = \frac{1+i\alpha}{1-i\alpha}\mathbf{y} \tag{B.6}$$

i.e. $\mathbf{x} = c\mathbf{y}, c \in \mathbb{C}, |c| = 1$. □

B.2 Appendix : Nature of Minima in $\mathbb{C}^n$

We first provide basics on mapping vectors in $\mathbb{C}^n$ to $\mathbb{R}^{2n}$.

B.2.1 Mapping from $\mathbb{C}^n$ to $\mathbb{R}^{2n}$

Let $a, b \in \mathbb{C}^n$. Define the operation $\mathcal{J} : \mathbb{C}^n \rightarrow \mathbb{R}^{2n}$ as follows,

$$\mathcal{J}(a+bi) = (a, b)$$

Define the conjugation operation $J : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n} \times \mathbb{R}^{2n}$ as:

$$J(\eta, \zeta) = (-\zeta, \eta) \qquad \forall \eta, \zeta \in \mathbb{R}^{2n}$$

Notice that J can be considered to be the matrix,

$$J = \begin{bmatrix} 0 & -I^{2n} \\ I^{2n} & 0 \end{bmatrix}$$

The outer product operation $[[.,.]] : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}^{n \times n}$ as follows:

$$[[\mathbf{u}, \mathbf{v}]] = \frac{1}{2} \left(\mathbf{u}\mathbf{v}^* + \mathbf{v}\mathbf{u}^* \right)$$

Let $A = \{\mathbf{a}_i\}_{i=1}^n, B = \{\mathbf{b}_i\}_{i=1}^n$ denote two orthonormal basis sets for $\mathbb{C}^n$. Define the linear operator $T_{A,B} : \mathbb{C}^n \to \mathbb{C}^n$ as follows,

$$T_{A,B}(\mathbf{u}) = \left(\sum_{i=1}^{n} \mathbf{b}_i \mathbf{a}_i^* \right) \mathbf{u} = \sum_{i=1}^{m} \langle \mathbf{u}, \mathbf{a}_i \rangle \mathbf{b}_i$$

Let $\zeta \in \mathbb{R}^{2n}$. Define $\tilde{T}$ to be the analogous operator of T in $\mathbb{R}^{2n}$ as $\tilde{T}_{A,B} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$,

$$\begin{aligned}
&\tilde{T}_{A,B}(\zeta) \\
&= \sum_{i=1}^{n} \langle \zeta, \mathcal{J}(\mathbf{a}_i) \rangle \mathcal{J}(\mathbf{b}_i) + \langle \zeta, \mathcal{J}(i\mathbf{a}_i) \rangle \mathcal{J}(i\mathbf{b}_i) \\
&= \left(\sum_{i=1}^{n} \mathcal{J}(\mathbf{b}_i) \mathcal{J}(\mathbf{a}_i)^\top + \mathcal{J}(i\mathbf{b}_i) \mathcal{J}(i\mathbf{a}_i)^\top \right) \zeta \\
&= \left(\sum_{i=1}^{n} \mathcal{J}(\mathbf{b}_i) \mathcal{J}(\mathbf{a}_i)^T + J\mathcal{J}(\mathbf{b}_i) \mathcal{J}(\mathbf{a}_i)^T J^T \right) \zeta \\
&= \left(\sum_{i=1}^{n} \mathcal{J}(\mathbf{b}_i) \mathcal{J}(\mathbf{a}_i)^T + J\mathcal{J}(\mathbf{b}_i) \mathcal{J}(\mathbf{a}_i)^T J^T \right) \zeta
\end{aligned}$$

Let $M = \sum_{i=1}^{n} \mathcal{J}(\mathbf{a}_i) \mathcal{J}(\mathbf{b}_i)^T$. Thus,

$$\tilde{T}_{A,B}(\zeta) = M\zeta + JMJ^T\zeta$$

Define the map $\tau : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{R}^{2n} \times \mathbb{R}^{2n}$ such that,

$$\tau(T_{A,B}) = \tilde{T}_{A,B}$$

Notice that the outer product $[[.,.]]$ under τ operation can be worked out to be Balan (2016),

$$\tau([[\mathbf{x}, \mathbf{y}]]) = [[\mathcal{J}(\mathbf{x}), \mathcal{J}(\mathbf{y})]] + [[\mathcal{J}(i\mathbf{x}), \mathcal{J}(i\mathbf{y})]]$$

Let $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$. We know $\mathbf{u} = \mathbf{u}_\mathbb{R} + i\mathbf{u}_\mathbb{C}, \mathbf{v} = \mathbf{v}_\mathbb{R} + i\mathbf{v}_\mathbb{C}$, $\mathbf{u}_\mathbb{R}, \mathbf{u}_\mathbb{C}, \mathbf{v}_\mathbb{R}, \mathbf{v}_\mathbb{C} \in \mathbb{R}^n$. The inner product operation $\langle ., . \rangle$ in $\mathbb{R}^{2n} \times \mathbb{R}^{2n}$ can be seen as,

$$\langle \mathcal{J}(\mathbf{u}), \mathcal{J}(\mathbf{v}) \rangle = \langle \mathbf{u}_\mathbb{R}, \mathbf{v}_\mathbb{R} \rangle + \langle \mathbf{u}_\mathbb{C}, \mathbf{v}_\mathbb{C} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle_\mathbb{R} \tag{B.7}$$

Note that $\forall A_k \in \mathcal{A}$, we can define orthogonal vectors $\{\mathbf{a}_i^k\}_{i=1}^n, \{\mathbf{b}_i^k\}_{i=1}^n$ such that $A_k = \sum_{i=1}^n \mathbf{a}_i^k (\mathbf{b}_i^k)^*$. Let $M_k = \sum_{i=1}^n \mathcal{J}(\mathbf{a}_i^k)(\mathcal{J}(\mathbf{b}_i^k))^T$. Define $\Phi_k \in \mathbb{R}^{2n \times 2n}$ as,

$$\Phi_k = \tau(A_k) = M_k + JM_kJ^T$$

Thus $\forall \zeta \in \mathbb{R}^{2n}$ we have,

$$\begin{aligned}\Phi_k\zeta &= M_k\zeta + JM_kJ^T\zeta \\ &= \sum_{i=1}^n \langle \zeta, \mathcal{J}(\mathbf{a}_i^k)\rangle \mathcal{J}(\mathbf{b}_i^k)^T + \langle \zeta, J\mathcal{J}(\mathbf{a}_i^k)\rangle (J\mathcal{J}(\mathbf{b}_i^k))^T\end{aligned}$$

Let $\{\mathbf{e}_i\}_{i=1}^n$ be the orthonormal basis for $\mathbb{C}^n$. Consider any rank-1 matrix $P \in \mathbb{C}^{n\times n}$. We can take, without loss of generality, $P = \mathbf{e}_1\mathbf{e}_1^H$, where $\mathbf{e}_1$ can be any unit norm vector in $\mathbb{C}^n$. We can easily show that

$$\tau(P) = \epsilon_1\epsilon_1^T + J\epsilon_1\epsilon_1^T J^T$$

where $\epsilon_1 = \mathcal{J}(\mathbf{e}_1)$.

Let $Q \in \mathbb{C}^{n\times n}$ be a rank-k matrix. Then,

$$\begin{aligned}Q &= \sum_{i=1}^k \mathbf{e}_i\mathbf{e}_i^H \quad \text{w.l.o.g} \\ \tau(Q) &= \sum_{i=1}^k \epsilon_i\epsilon_i^T + J\epsilon_i\epsilon_i^T J^T, \quad \epsilon_i = \mathcal{J}(\mathbf{e}_i)\end{aligned}$$

Using (B.7), notice that the following conditions on the orthonormal basis hold,

$$\begin{aligned}\langle J\epsilon_l, J\epsilon_k\rangle &= \langle \epsilon_l, \epsilon_k\rangle = \langle \mathbf{e}_l, \mathbf{e}_k\rangle_{\mathbb{R}} = \delta_{l,k} \\ \langle J\epsilon_l, \epsilon_k\rangle &= \langle i\mathbf{e}_l, \mathbf{e}_k\rangle_{\mathbb{R}} = 0\end{aligned}$$

Note that $\{\epsilon_i, J\epsilon_i\}_1^k$ spans $\tau(Q)$ hence making the rank of $\tau(Q)$ as 2k.

In order to observe properties of the mapping $\mathcal{M}_{\mathcal{A}}$ in the real vector space $\mathbb{R}^{2n}$, we need to find the analog of $\langle A_i, \mathbf{x}\mathbf{x}^H\rangle$ in $\mathbb{R}^{2n}$.

Towards this end, we prove a relation between $tr\{\tau(T), \tau(S)\}$ and $tr\{TS\}$. Notice that since $\{\epsilon_i, J\epsilon_i\}_{i=1}^n$ is orthonormal basis for $\mathbb{R}^{2n}$, the following holds,

$$\begin{aligned}&tr\{\tau(T), \tau(S)\} \\ &= \sum_{i=1}^n \langle \tau(S)\epsilon_i, \tau(T)\epsilon_i\rangle + \langle \tau(S)J\epsilon_i, \tau(T)J\epsilon_i\rangle \\ &= \sum_{i=1}^n \langle S\mathbf{e}_i, T\mathbf{e}_i\rangle + \langle Si\mathbf{e}_i, Ti\mathbf{e}_i\rangle \\ &= 2\sum_{i=1}^n \langle S\mathbf{e}_i, T\mathbf{e}_i\rangle \\ &= 2\langle T, S\rangle \\ &= 2tr\{TS\}\end{aligned} \tag{B.8}$$

B.2.2 Nature of Global Minima in $\mathbb{C}^n$

We prove the following theorem to provide insights into the properties of the mapping $\mathcal{M}_{\mathcal{A}}$.

Theorem B.2.1. *The following statements are equivalent:*

1. *There exists constants $\alpha, \beta > 0$ such that $\forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$, we have that,*

$$b_0 \|[[\boldsymbol{u}, \boldsymbol{v}]]\|_1^2 \geq \sum_{i=1}^{m} |\langle A_k, [[\boldsymbol{u}, \boldsymbol{v}]] \rangle|^2 \geq a_0 \|[[\boldsymbol{u}, \boldsymbol{v}]]\|_1^2 \tag{B.9}$$

2. *$\forall \eta \in \mathbb{R}^{2n} \neq 0$, $\exists$ $c_0, d_0 > 0$ such that,*

$$d_0 \|\eta\|^2 P_{J\eta}^{\perp} \succeq R(\eta) \succeq c_0 \|\eta\|^2 P_{J\eta}^{\perp} \tag{B.10}$$

Where,

$$P_{J\eta}^{\perp} = I - \frac{1}{\|\eta\|^2} J\eta\eta^T J^T$$

Next, we prove (2)↔1
Note that,

$$\begin{aligned} \langle \tau(A_i), \tau([[\mathbf{u}, \mathbf{v}]]) \rangle &= \langle M_i + JM_iJ^T, [[\zeta, \eta]]) \rangle \\ &= [\zeta^T M_i \eta + \zeta^T JM_iJ^T \eta] \\ &= \zeta^T \Phi_i \eta \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^{m} \left(\langle \tau(A_i), \tau([[\mathbf{u}, \mathbf{v}]]) \rangle \right)^2 &= \sum_{i=1}^{m} (\zeta^T \Phi_i \eta)^2 \\ &= \sum_{i=1}^{m} \zeta^T \left(\Phi_i \eta\eta^T \Phi_i^T \right) \zeta \end{aligned}$$

Let us defined $R(\eta) = \sum_{i=1}^{m} \Phi_i \eta\eta^T \Phi_i^T$. Thus we can rewrite,

$$\sum_{i=1}^{m} \left(\langle \tau(A_i), \tau([[\mathbf{u}, \mathbf{v}]]) \rangle \right)^2 = \zeta^T R(\eta) \zeta$$

From Balan (2016), note that,

$$\begin{aligned} \|[[\mathbf{u}, \mathbf{v}]]\|_F^2 &= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (imag(\langle \mathbf{u}, \mathbf{v} \rangle))^2 \\ &= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (real(\langle i\mathbf{u}, \mathbf{v} \rangle))^2 \\ &= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - \langle i\mathbf{u}, \mathbf{v} \rangle_{\mathbb{R}}^2 \\ &= \|\zeta\|^2 \|\eta\|^2 - (\langle J\zeta, \eta \rangle)^2 \\ &= \zeta^T \left(\|\eta\|^2 I - J\eta\eta^T J^T \right) \zeta \end{aligned}$$

where $\mathcal{J}(\mathbf{u}) = \zeta, \mathcal{J}(\mathbf{v}) = \eta$.

Using the results from 1 and equation (B.8), We can write,

$$b_0\|\eta\|^2 P_\perp \succeq R(\eta) \succeq a_0\|\eta\|^2 P_\perp$$

where $P_\perp = \left(I - \frac{1}{\|\eta\|^2} J\eta\eta^H J^H\right)$.

Hence $1 \Leftrightarrow 2$ holds with the same constants $c_0 = a_0$ and $d_0 = b_0$.

$\forall \eta \in \mathbb{R}^{2n} \neq 0$, $\exists\ c_0, d_0 > 0$ such that,

$$d_0\|\eta\|^2 P^\perp_{J\eta} \succeq R(\eta) \succeq c_0\|\eta\|^2 P^\perp_{J\eta}$$

where,

$$P^\perp_{J\eta} = I - \frac{1}{\|\eta\|^2} J\eta\eta^T J^T$$

Lemma B.2.2. *for any $\eta \in \mathbb{R}^{2n} \neq 0$, we have that rank(R(η)) = 2n-1 .*

Proof. We know,

$$P_\perp = \left(I - \frac{1}{\|\eta\|^2} J\eta\eta^H J^H\right)$$

Consider $\zeta, \gamma \in \mathbb{R}^{2n}$ such that $\zeta = cJ\eta + \gamma$ for some constant $c \in \mathbb{R}$, then

$$\begin{aligned}
&\left\langle \left(I - \frac{1}{\|\eta\|^2} J\eta\eta^T J^T\right)\zeta, \zeta \right\rangle \\
&= \|\zeta\|^2 - \frac{1}{\|\eta\|^2}\langle\zeta, J\eta\rangle\langle J\eta, \zeta\rangle \\
&= \|cJn + \lambda\|^2 - \frac{1}{\|\eta\|^2}\left(\langle cJ\eta + \lambda, J\eta\rangle\right)^2 \\
&= \|cJn + \lambda\|^2 - \frac{1}{\|\eta\|^2}\left(c\|J\eta\|^2 + \langle\lambda, J\eta\rangle\right)^2 \\
&= c^2\|Jn\|^2 + 2c\langle J\eta, \lambda\rangle + \|\lambda\|^2 \\
&- \frac{1}{\|\eta\|^2}\left(c^2\|J\eta\|^4 + 2c\|J\eta\|^2\langle\lambda, J\eta\rangle + \langle\lambda, J\eta\rangle^2\right)
\end{aligned}$$

Note that $\|J\eta\| = \|\eta\|$. Hence we can simplify the expression as,

$$\left\langle \left(I - \frac{1}{\|\eta\|^2} J\eta\eta^T J^T\right)\zeta, \zeta \right\rangle = \|\lambda\|^2 - \frac{\langle\lambda, J\eta\rangle^2}{\|\eta\|^2}$$

As can be noted from above, that the values of λ at which the above expression vanishes is $\lambda = 0$.

Hence we can conclude that the rank of $P^\perp_{J\eta}$ is 2n-1. And taking into account equation (B.10), we have our required result. □

Lemma B.2.2 shows that the value of the bounds on the function $R(\eta_0)$ will not change at $R(\eta_0 + J\eta)$.

Notice that the operation J in the $\mathbb{R}^{2n}$ is analogue to shifting the phase by multiplying complex vector by i.

For the vectors $\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ satisfying equation (B.1), the analogous distance metric satisfies the property,

$$\begin{aligned}\tau(d(\mathbf{x},\mathbf{y})) &= \tau(\|\mathbf{x}\mathbf{x}^H - \mathbf{y}\mathbf{y}^H\|_F) \\ &= \tau([[\mathbf{u},\mathbf{v}]]) \\ &= \zeta^T P^{\perp}_{J\eta}\zeta\end{aligned}$$

and hence, considering that rank($P^{\perp}_{J\eta}$) =2n-1, there can be only one direction for the distance metric where the value doesn't change. This is the direction of the phase shift of any of the vectors $\mathbf{x}, \mathbf{y}$.

B.3 Appendix : Non-convex Landscape for Deterministic $\mathcal{A}$

B.3.1 Supporting Results

Throughout the rest of the appendix, define $\Delta = \mathbf{x} - e^{i\phi}\mathbf{z}$ such that $\phi = \underset{\theta\in[0,2\pi]}{\arg\min}\|\mathbf{x} - e^{i\theta}\mathbf{z}\|_2$ for any $\mathbf{x}, \mathbf{z} \in \mathbf{C}^n$.

Lemma B.3.1. *For any* $\mathbf{x} \in \mathbf{C}^n$,

$$\|\mathbf{x}\mathbf{x}^*\|_F^2 = \|\mathbf{x}\|_2^4$$

Proof.

$$\begin{aligned}&\|\mathbf{x}\mathbf{x}^*\|_F^2 \\ &= \sum_{i,j=1}^{n} |x_i\bar{x}_j|^2 = \sum_{i,j=1}^{n} (x_i\bar{x}_j)^*(x_i\bar{x}_j) = \sum_{i,j=1}^{n} x_j\bar{x}_i x_i\bar{x}_j \\ &= \sum_{i,j=1}^{n} |x_j|^2|x_i|^2 = (\sum_{i=1}^{n} |x_i|^2)^2 = \|\mathbf{x}\|_2^4\end{aligned}$$

□

Lemma B.3.2. *The vectors* $\mathbf{x} - e^{i\phi}\mathbf{z}$ *and* $\mathbf{x} + e^{i\phi}\mathbf{z}$ *are such that* $Im(\langle \mathbf{x} - e^{i\phi}\mathbf{z}, \mathbf{x} + e^{i\phi}\mathbf{z}\rangle) = 0$, *where* $\phi = \underset{\theta\in[0,2\pi]}{\arg\min}\|\mathbf{x} - e^{i\theta}\mathbf{z}\|_2$

Proof. We know from Lemma 3.7 Balan (2016) that $\forall \mathbf{x}, \mathbf{z} \in \mathbb{C}^n$ $\exists \mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ such that,

$$\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^* = \mathbf{u}\mathbf{v}^* + \mathbf{v}\mathbf{u}^* = [[\mathbf{u},\mathbf{v}]]$$

It can be easily verified that few such pairs $\mathbf{u}, \mathbf{v}$ are given by,

$$\mathbf{u} = \mathbf{x} - e^{i\theta}\mathbf{z}, \quad \mathbf{v} = \mathbf{x} + e^{i\theta}\mathbf{z}, \quad \forall\theta \in [0, 2\pi]$$

Next, we focus on $\langle \mathbf{u}, \mathbf{v} \rangle$. We argue that $\exists \theta$ such that $\text{Im}(\langle (\mathbf{x} - e^{i\theta}\mathbf{z}), (\mathbf{x} + e^{i\theta}\mathbf{z}) \rangle) = 0$ To this end consider the following,

$$\begin{aligned}
\langle \mathbf{u}, \mathbf{v} \rangle &= \mathbf{u}^T \bar{\mathbf{v}} \\
&= \left(\mathbf{x} - e^{i\theta}\mathbf{z}\right)^T \overline{(\mathbf{x} + e^{i\theta}\mathbf{z})} \\
&= \left(\mathbf{x} - e^{i\theta}\mathbf{z}\right)^T \left(\bar{\mathbf{x}} + e^{-i\theta}\bar{\mathbf{z}}\right) \\
&= \langle \mathbf{x}, \mathbf{x} \rangle - e^{i\theta}\langle \mathbf{z}, \mathbf{x} \rangle + e^{-i\theta}\langle \mathbf{x}, \mathbf{z} \rangle - \langle \mathbf{z}, \mathbf{z} \rangle \\
&= \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{z}, \mathbf{z} \rangle - 2i\text{Im}(e^{i\theta}\langle \mathbf{z}, \mathbf{x} \rangle)
\end{aligned}$$

$\text{Im}(e^{i\theta}\langle \mathbf{z}, \mathbf{x} \rangle)$ can only vanish if $\theta = \omega$ where $\omega \in [0, 2\pi]$ is the angle between the two vectors, i.e. ω is such that $\langle \mathbf{x}, \mathbf{z} \rangle = e^{i\omega}\|\mathbf{x}\|\|\mathbf{z}\|$. Next we prove that $\omega = \phi$ where,

$$\phi = \underset{\theta \in [0, 2\pi]}{\arg\min} \|\mathbf{x} - e^{i\theta}\mathbf{z}\|_2$$

Consider the following argument,

$$\begin{aligned}
\underset{\theta \in [0, 2\pi]}{\arg\min} \|\mathbf{x} - e^{i\theta}\mathbf{z}\|_2^2 &= \underset{\theta \in [0, 2\pi]}{\arg\min} \left(\mathbf{x} - e^{i\theta}\mathbf{z}\right)^* \left(\mathbf{x} - e^{i\theta}\mathbf{z}\right) \\
&= \underset{\theta \in [0, 2\pi]}{\arg\min} \|\mathbf{x}\|^2 + \|\mathbf{z}\|^2 - e^{-i\theta}\mathbf{z}^*\mathbf{x} - e^{i\theta}\mathbf{x}^*\mathbf{z} \\
&= \underset{\theta \in [0, 2\pi]}{\arg\min} \|\mathbf{x}\|^2 + \|\mathbf{z}\|^2 - 2Re(e^{-i\theta}\langle \mathbf{x}, \mathbf{z} \rangle) \\
&= \|\mathbf{x}\|^2 + \|\mathbf{z}\|^2 - 2\underset{\theta \in [0, 2\pi]}{\arg\max} Re(e^{-i\theta}\langle \mathbf{x}, \mathbf{z} \rangle)
\end{aligned}$$

It can be seen easily that the $\underset{\theta \in [0, 2\pi]}{\arg\max} Re(e^{-i\theta}\langle \mathbf{x}, \mathbf{z} \rangle)$ is achieved when $\theta = \omega$. Thus we have proved that $\phi = \omega$. □

Lemma B.3.3. *Let* $\mathbf{x}, \mathbf{z} \in \mathbb{C}^n$. *Then,*

$$\|(\mathbf{x} - e^{i\phi}\mathbf{z})(\mathbf{x} - e^{i\phi}\mathbf{z})^*\|_F^2 \leq 2\|\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^*\|$$

where $\phi = \underset{\theta \in [0, 2\pi]}{\arg\min} \|\mathbf{x} - e^{i\theta}\mathbf{z}\|$

Proof. Note,

$$\begin{aligned}
\underset{\theta \in [0, 2\pi]}{\arg\min} \|\mathbf{x} - e^{i\theta}\mathbf{z}\|^2 &= \underset{\theta \in [0, 2\pi]}{\arg\min} \left(\mathbf{x} - e^{i\theta}\mathbf{z}\right)^* \left(\mathbf{x} - e^{i\theta}\mathbf{z}\right) \\
&= \underset{\theta \in [0, 2\pi]}{\arg\min} \|\mathbf{x}\|^2 + \|\mathbf{z}\|^2 - e^{-i\theta}\mathbf{z}^*\mathbf{x} - e^{i\theta}\mathbf{x}^*\mathbf{z} \\
&= \underset{\theta \in [0, 2\pi]}{\arg\min} \|\mathbf{x}\|^2 + \|\mathbf{z}\|^2 - 2Re(\langle \mathbf{x}, e^{i\theta}\mathbf{z} \rangle)
\end{aligned}$$

The minimum can only be achieved at a point where $Re(\mathbf{x}^*(e^{i\phi}\mathbf{z})) \geq 0$. Further notice that the following relation holds,

$$\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^* + \Delta\Delta^* = \mathbf{x}\Delta^* + \Delta\mathbf{x}^* \tag{B.11}$$

Hence we can see that,

$$\|\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^*\|_F^2 = \|\mathbf{x}\Delta^* + \Delta\mathbf{x}^* - \Delta\Delta^*\|_F^2$$

We know that for any matrix A, $\|A\|_F^2 = \mathrm{Tr}(A^H A)$,

$$\begin{aligned}
&\|\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^*\|_F^2 \\
&= Tr\left((\mathbf{x}\Delta^* + \Delta\mathbf{x}^* - \Delta\Delta^*)^*(\mathbf{x}\Delta^* + \Delta\mathbf{x}^* - \Delta\Delta^*)\right) \\
&= \|\mathbf{x}\Delta^*\|_F^2 + (\langle \mathbf{x}, \Delta\rangle)^2 + (\langle \Delta, \mathbf{x}\rangle)^2 + \|\Delta\mathbf{x}^*\|_F^2 \\
&- 2\langle \mathbf{x}, \Delta\rangle\|\Delta\|_F^2 - 2\langle \Delta, \mathbf{x}\rangle\|\Delta\|_F^2 + \|\Delta\Delta^*\|_F^2
\end{aligned}$$

Note that,

$$\begin{aligned}
(\langle \mathbf{x}, \Delta\rangle)^2 + (\langle \Delta, \mathbf{x}\rangle)^2 &= (\langle \mathbf{x}, \Delta\rangle)^2 + \overline{(\langle \mathbf{x}, \Delta\rangle)^2} \\
&= 2Re(\langle \mathbf{x}, \Delta\rangle)^2 \\
\langle \mathbf{x}, \Delta\rangle\|\Delta\|_F^2 + \langle \Delta, \mathbf{x}\rangle\|\Delta\|_F^2 &= \langle \mathbf{x}, \Delta\rangle\|\Delta\|_F^2 + \overline{\langle \mathbf{x}, \Delta\rangle}\|\Delta\|_F^2 \\
&= 2Re(\langle \mathbf{x}, \Delta\rangle)\|\Delta\|_F^2
\end{aligned}$$

Thus we can conclude,

$$\begin{aligned}
&\|\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^*\|_F^2 \\
&= 2\|\langle \mathbf{x}, \Delta\rangle\|_F^2 + 2Re((\langle \mathbf{x}, \Delta\rangle)^2) \\
&- 4Re(\langle \mathbf{x}, \Delta\rangle)\|\Delta\|_F^2 + \|\Delta\Delta^*\|_F^2 \\
&= 2\mathbf{x}^*\mathbf{x}\Delta^*\Delta + 2Re((\langle \mathbf{x}, \Delta\rangle)^2) \\
&- 4Re(\mathbf{x}^*\Delta\Delta^*\Delta) + \|\Delta\Delta^*\|_F^2
\end{aligned}$$

Since $\mathbf{x}^*\mathbf{x}\Delta^*\Delta = \|\langle \mathbf{x}, \Delta\rangle\|_F^2$, its a real value.

$$\begin{aligned}
&\|\mathbf{x}\mathbf{x}^* - \mathbf{x}^*(\mathbf{x}^*)^*\|_F^2 \\
&= 2\mathbf{x}^*\left(\mathbf{x} - \Delta\right)\Delta^*\Delta + 2Re((\langle \mathbf{x}, \Delta\rangle)^2) \\
&- 2Re(\mathbf{x}^*\Delta\Delta^*\Delta) + \|\Delta\Delta^*\|_F^2 \\
&= 2\mathbf{x}^*\left(\mathbf{x} - \Delta\right)\Delta^*\Delta + (\langle \mathbf{x}, \Delta\rangle)^2 + (\langle \Delta, \mathbf{x}\rangle)^2 \\
&- \langle \mathbf{x}, \Delta\rangle\|\Delta\|_F^2 - \langle \Delta, \mathbf{x}\rangle\|\Delta\|_F^2 + \|\Delta\Delta^*\|_F^2 \\
&= 2\mathbf{x}^*\left(\mathbf{x} - \Delta\right)\Delta^*\Delta + \left(\langle \mathbf{x}, \Delta\rangle - \frac{1}{2}\langle \Delta, \Delta\rangle\right)^2 \\
&+ \left((\langle \Delta, \mathbf{x}\rangle - \frac{1}{2}\langle \Delta, \Delta\rangle\right)^2 + \frac{1}{2}\|\Delta\Delta^*\|_F^2 \\
&= 2\mathbf{x}^*\left(\mathbf{x} - \Delta\right)\Delta^*\Delta + 2Re\left((\langle \Delta, \mathbf{x} - \frac{1}{2}\Delta\rangle\right)^2 \\
&+ \frac{1}{2}\|\Delta\Delta^*\|_F^2 \\
&= 2\mathbf{x}^* e^{i\phi}\mathbf{z}\Delta^*\Delta + \frac{1}{2}Re\left((\langle \Delta, \mathbf{x} + e^{i\phi}\mathbf{z}\rangle\right)^2 + \frac{1}{2}\|\Delta\Delta^*\|_F^2
\end{aligned}$$

From Lemma B.3.2, it can be seen that $Im\left(\langle \Delta, \mathbf{x} + e^{i\phi}\mathbf{z}\rangle\right) = 0$. Hence we can say,

$$\|\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^*\|_F^2 \geq \frac{1}{2}\|\Delta^*\Delta\|_F^2$$

□

B.3.2 Wirtinger Calculus

We use standard arguments from wirtinger calculus Wang and Xu (2017) to prove results Theorem 3.6.7. The basic intuition is to look at the ℓ_2-loss function f (F2) as function of two real variables in $\mathbb{R}^n$ rather instead of single complex variable in $\mathbb{C}^n$. This workaround is required for the analysis of real function of complex variables because of notions of complex differentiability and conclusions from Cauchy-Reimann equations Shunmugaraj (????). Hence we map the function $f : \mathbb{C}^n \to \mathbb{R}$ to $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and instead of analysing the properties of $\nabla^2 f$, we analyse the properties of $\nabla^2 g$.

We first introduce the mapped function g and the corresponding expressions for ∇g and $\nabla^2 g$

$$\begin{aligned} f(\mathbf{x}) = g(\mathbf{x}, \bar{\mathbf{x}}) &= \frac{1}{m}\sum_{i=1}^{m} g_i(\mathbf{x}, \bar{\mathbf{x}}) \\ &= \frac{1}{m}\sum_{i=1}^{n} |\bar{\mathbf{x}}^\top A_i \mathbf{x} - c_i|^2 \end{aligned}$$

For the gradient ∇g we have,

$$\nabla g(\mathbf{x}, \bar{\mathbf{x}}) = \frac{1}{m}\sum_{i=1}^{n} \begin{bmatrix} (\bar{\mathbf{x}}^\top A_i \mathbf{x} - c_i) A_i \mathbf{x} \\ (\bar{\mathbf{x}}^\top A_i \mathbf{x} - c_i) A_i \bar{\mathbf{x}} \end{bmatrix}$$

For the hessian $\nabla^2 g$, we have

$$\nabla^2 g(\mathbf{x}, \bar{\mathbf{x}}) = \frac{1}{m}\sum_{i=1}^{m} \begin{bmatrix} (2\bar{\mathbf{x}}^\top A_i \mathbf{x} - c_i) A_i & (A_i \mathbf{x})(A_i \mathbf{x})^\top \\ (A_i \bar{\mathbf{x}})(A_i \bar{\mathbf{x}})^\top & (2\bar{\mathbf{x}}^\top A_i \mathbf{x} - c_i) A_i \end{bmatrix}$$

The following can be verified easily,

$$\begin{aligned} &\langle \nabla g(\mathbf{x}), \begin{bmatrix} \Delta \\ \bar{\Delta} \end{bmatrix} \rangle \\ &= \frac{1}{m}\sum_{i=1}^{m} \langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top \rangle \langle A_i, \mathbf{x}\bar{\Delta}^\top + \Delta \bar{\mathbf{x}}^\top \rangle \\ &= \frac{1}{m}\sum_{i=1}^{m} \langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top \rangle \langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}}^*)^\top + \Delta \bar{\Delta}^\top \rangle \end{aligned} \tag{B.12}$$

$$
\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g(\mathbf{x},\bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\\
&= \frac{1}{m}\sum_{i=1}^{m} \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_i(\mathbf{x},\bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\\
&= \frac{1}{m}\sum_{i=1}^{m} (2\mathbf{x}^\top A_i\bar{\mathbf{x}} - b_i)(\bar{\Delta}^\top A_i\Delta + \Delta^\top A_i\bar{\Delta})\\
&+ \left((\Delta^\top A_i\bar{\mathbf{x}})^2 + (\bar{\Delta}A_i\mathbf{x})^2\right)
\end{aligned}
$$

Theorem B.3.4. *Let $\mathcal{A}$ be a set of measurement matrices which satisfy (α,β)-stability with $2\beta < 3\alpha$ and κ-phase-discriminating with $\kappa \geq 0$. Let the scalar vector $\mathbf{c}$ be generated by quadratic measurements of an unknown vector $\mathbf{z}$ characterizing the measurements used in the objective function f of problem (P2) . Then the following statements holds:*

1) The function f is strict saddle, and

2) Every local minimum $\mathbf{w}$ of f satisfies $d(\mathbf{w},\mathbf{z}) = 0$

Proof. Notice that,

$$
\begin{aligned}
&(\Delta A\bar{\mathbf{x}})^2 + (\bar{\Delta}^\top A\mathbf{x})^2\\
&= \left(\langle A, \mathbf{x}\bar{\Delta}^\top + \Delta\bar{\mathbf{x}}^\top\rangle\right)^2 - 2(\Delta^\top A\bar{\mathbf{x}})(\bar{\Delta}^\top A\mathbf{x})\\
&= \left(\langle A, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top + \Delta\bar{\Delta}^\top\rangle\right)^2 - 2(\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)
\end{aligned}
$$

Using (B.11), we can reorganize,

$$
\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_i(\mathbf{x},\bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\\
&= \langle A_i, 2\mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, 2\Delta\bar{\Delta}^\top\rangle + \left(\langle A, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top + \Delta\bar{\Delta}^\top\rangle\right)^2\\
&- 2(\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)\\
&= 2\left(\langle A_i, 2\mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle\right)\\
&+ \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle + \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle\\
&+ \langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top + \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\\
&- 2(\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)
\end{aligned}
$$

Adding and subtracting $2\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle$, reorganizing,

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_i(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\\
&= 2\langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top\rangle\\
&+ 2\left(\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top + \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\right)\\
&- 2\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\\
&+ \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\\
&+ \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle - 2(\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)\\
&+ \langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top + \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle
\end{aligned}$$

Adding and subtracting $\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle$, reorganizing,

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_i(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\\
&= 2\langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top\rangle - 2(\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)\\
&- 3\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\\
&+ \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle\\
&+ 4\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top + \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle
\end{aligned}$$

Using equation (B.12),

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_i(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\\
&= 2\langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top\rangle - 2(\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)\\
&- 3\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\\
&+ \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle + 4\langle \nabla g_i(\mathbf{x}, \bar{\mathbf{x}}), \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\rangle
\end{aligned}$$

Overall, we can conclude that,

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix} = \frac{1}{m}\sum_{i=1}^m \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_i(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\\
&= \frac{2}{m}\sum_{i=1}^m \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top\rangle - (\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)\\
&- \frac{3}{m}\sum_{i=1}^m \langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\\
&+ \frac{1}{m}\sum_{i=1}^m \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle + \frac{4}{m}\sum_{i=1}^m \langle \nabla g_i(\mathbf{x}, \bar{\mathbf{x}}), \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\rangle
\end{aligned}$$

Using the definition of (α, β)-stable and κ-cross stable, we can conclude that for every critical point ($\sum_{i=1}^{m} \nabla g_i(\mathbf{x}) = 0$),

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix} \\
&\leq -\kappa \|\Delta \mathbf{x}^T - \mathbf{x}\Delta^T\|_F^2 + \beta \|\Delta\Delta^*\|_F^2 - 3\alpha \|\mathbf{x}\mathbf{x}^* - \mathbf{z}(\mathbf{z})^*\|_F^2 \\
&\leq -\kappa \|\Delta \mathbf{x}^T - \mathbf{x}\Delta^T\|_F^2 + 2\beta \|\mathbf{x}\mathbf{x}^* - \mathbf{z}(\mathbf{z})^*\|_F^2 \\
&- 3\alpha \|\mathbf{x}\mathbf{x}^* - \mathbf{z}(\mathbf{z})^*\|_F^2 \\
&\leq (2\beta - 3\alpha) \|\mathbf{x}\mathbf{x}^* - \mathbf{z}(\mathbf{z})^*\|_F^2 - \kappa \|\Delta \mathbf{x}^T - \mathbf{x}\Delta^T\|_F^2
\end{aligned}$$

as stated in the statement of the Theorem, since $2\beta < 3\alpha$ and $\kappa > 0$, we can conclude that the function $f(\mathbf{x}) = g(\mathbf{x}, \overline{\mathbf{x}})$ satisfies at-least one of the following is true,

- $\|\nabla g(\mathbf{x})\| > 0$
- $\forall \mathbf{x} \in \mathbb{C}^n$ such that $d(\mathbf{x}, \mathbf{z}) \neq 0$, the following holds,

$$\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix} < 0$$

- $d(\mathbf{x}, \mathbf{z}) = 0$

Thus the function $f(\mathbf{x})$ is strict saddle.

Following up on equation (B.28), the only possible way that the hessain $\nabla^2 g(\mathbf{x}, \overline{\mathbf{x}}) \succeq 0$ is if $\|\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^*\|_F = 0$ and $\|\Delta \mathbf{x}^T - \mathbf{x}\Delta^T\|_F = 0$. Hence we can conclude that all local minimas, i.e. all $\mathbf{w}$ such that $\nabla^2 g(\mathbf{w}, \overline{\mathbf{w}}) \succeq 0$ has to satisfy $\|\mathbf{w}\mathbf{w}^* - \mathbf{z}\mathbf{z}^*\| = 0$, $\|\Delta \mathbf{x}^T - \mathbf{x}\Delta^T\|_F = 0$ and hence satisfies $\mathbf{z} \sim \mathbf{w}$ which makes $\mathbf{w}$ the solution of the problem (P1).

Following up on equation (B.28), the only possible way that the hessain $\nabla^2 g(\mathbf{x}, \overline{\mathbf{x}}) \succeq 0$ is if $\|\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^*\| = 0$. Hence we can conclude that all local minimas, i.e. all $\mathbf{w}$ such that $\nabla^2 g(\mathbf{w}, \overline{\mathbf{w}}) \succeq 0$ has to satisfy $\|\mathbf{w}\mathbf{w}^* - \mathbf{z}\mathbf{z}^*\| = 0$ and hence satisfies $\mathbf{z} \sim \mathbf{w}$ which makes $\mathbf{w}$ the solution of the problem (P1). □

B.4 Appendix : Properties of Gaussian Random $\mathcal{A}$

Lemma B.4.1. *Let $\mathcal{A} = \{A_i\}_{i=1}^m$ be a set of complex Hermitian Gaussian random matrices for the measurement model given by* (P1)*. Then, given $\epsilon > 0$ and vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, there are constants $c, d > 0$ such that*

$$\mathbb{P}\left(\left|\sum_{i=1}^{m} \frac{1}{m} |\langle A_i, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2 - d(\mathbf{x}, \mathbf{y})^2\right| \geq \epsilon d(\mathbf{x}, \mathbf{y})^2\right) \leq d e^{-cm\epsilon}.$$

Proof. A matrix $A \in \mathbb{C}^{n \times n}$ is a complex Hermitian Gaussian random matrix, if,

1. $\forall\ i$, $a_{ii} \sim \mathcal{N}(0, \sigma^2)$.

2. $\forall\ i, j,\ i \neq j,\ a_{ij} \sim \mathcal{N}(0, \frac{\sigma^2}{2}) + i\mathcal{N}(0, \frac{\sigma^2}{2})$.

Let $\{A_d\}_{d=1}^m$ be a set of complex Hermitian Gaussian random matrices. Define the random variable Y,

$$\begin{aligned} Y &= \frac{1}{m}\sum_{d=1}^{m} |\langle A_d, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^* \rangle|^2 \\ &= \frac{1}{m}\sum_{d=1}^{m} (\langle A_d, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^* \rangle)(\langle A_d, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^* \rangle) \\ &= \frac{1}{m}\sum_{d=1}^{m} \left(\sum_{ij} a_{ij}(x_i\bar{x}_j - y_i\bar{y}_j)\right)\left(\sum_{kl} a_{kl}(x_k\bar{x}_l - y_k\bar{y}_l)\right) \end{aligned}$$

Expectation of Y can be evaluated as,

$$\begin{aligned} &\mathbb{E}[Y] \\ &= \mathbb{E}\left(\frac{1}{m}\sum_{d=1}^{m} \left(\sum_{ij} a_{ij}(x_i\bar{x}_j - y_i\bar{y}_j)\right)\left(\sum_{kl} a_{kl}(x_k\bar{x}_l - y_k\bar{y}_l)\right)\right) \end{aligned} \tag{B.13}$$

For every matrix A_d, we can split the entire summation (B.13) into the following 4 sets:

1. $B := \{(i,j,k,l) | i = j = k = l\}$
2. $C := \{(i,j,k,l) | i = l, j = k\} \cap A^C$
3. $D := \{(i,j,k,l) | i = k, j = l\} \cap A^C$
4. $E := \{(i,j,k,l)\} \cap A^C \cap B^C \cap C^C$

Calculating the expectation of the sum of the elements in each individual sets:

1. Set B,

$$\begin{aligned} &\mathbb{E}\left(\sum_{(i,j,k,l)\in B} a_{ij}a_{kl}(x_i\bar{x}_j - y_i\bar{y}_j)(x_k\bar{x}_l - y_k\bar{y}_l)\right) \\ &= \mathbb{E}\left(\sum_{i=1}^{n} |a_{ii}|^2(|x_i|^2 - |y_i|^2)^2\right) \\ &= \sigma^2 \sum_{i=1}^{n} |x_i|^4 + |y_i|^4 - 2|x_i|^2|y_i|^2 \end{aligned}$$

2. Set C, note that for every Hermitian matrix A_d, $a_{ij} = \bar{a}_{ji}$

$$\begin{aligned}
&\mathbb{E}\left(\sum_{(i,j,k,l)\in C} a_{ij}a_{kl}(x_i\bar{x}_j - y_i\bar{y}_j)(x_k\bar{x}_l - y_k\bar{y}_l)\right) \\
&= \mathbb{E}\left(\sum_{i,j=1,i\neq j}^{n} |a_{ij}|^2(|x_i|^2|x_j|^2 - y_i\bar{y}_j x_j\bar{x}_i\right. \\
&\left.-x_i\bar{x}_j y_j\bar{y}_i + |y_i|^2|y_j|^2)\right) \\
&= \sigma^2 \sum_{i,j=1,i\neq j}^{n} (|x_i|^2|x_j|^2 - y_i\bar{y}_j x_j\bar{x}_i \\
&- x_i\bar{x}_j y_j\bar{y}_i + |y_i|^2|y_j|^2)
\end{aligned}$$

3. Set D,

$$\begin{aligned}
&\mathbb{E}\left(\sum_{(i,j,k,l)\in D} a_{ij}a_{kl}(x_i\bar{x}_j - y_i\bar{y}_j)(x_k\bar{x}_l - y_k\bar{y}_l)\right) \\
&= \mathbb{E}\left(\sum_{ij}(a_{ij})^2(x_i\bar{x}_j - y_i\bar{y}_j)^2)\right) = 0
\end{aligned}$$

Notice that $\forall i, j$

$$(a_{ij})^2 \ = ((a^r_{ij})^2 - (a^i_{ij})^2 + ia^r_{ij}a^i_{ij})$$

Thus,

$$\mathbb{E}\left[(a_{ij})^2\right] = \mathbb{E}\left[(a_{ij})^2\right] = \mathbb{E}\left[((a^r_{ij})^2 - (a^i_{ij})^2 + ia^r_{ij}a^i_{ij})\right]$$

Since both the real and imaginary parts are independent, we can conclude, $\mathbb{E}\left[((a^r_{ij})^2 - (a^i_{ij})^2 + ia^r_{ij}a^i_{ij})\right] = 0$

4. Set E,

All elements a_{ij}, a_{kl} are independent in $(i, j, k, l) \in E$,

$$\mathbb{E}\left(\sum_{(i,j,k,l)\in E} a_{ij}a_{kl}(x_i\bar{x}_j - y_i\bar{y}_j)(x_k\bar{x}_l - y_k\bar{y}_l)\right) = 0$$

In conclusion,

$$\begin{aligned}
&\mathbb{E}[Y] \\
&= \sigma^2 \left(\left(\sum_{i=1}^{n} |x_i|^2 \right)^2 + \left(\sum_{i=1}^{n} |y_i|^2 \right)^2 \right. \\
&\left. -2 \sum_{i=1}^{n} |x_i|^2 |y_i|^2 - \sum_{i,j,i \neq j} y_i \bar{y}_j x_j \bar{x}_i - \sum_{i,j,i \neq j} y_j \bar{y}_i x_i \bar{x}_j \right) \\
&= \sigma^2 \left[\|\mathbf{x}\|_2^4 + \|\mathbf{y}\|_2^4 - |\langle \mathbf{x}, \mathbf{y} \rangle|^2 \right]
\end{aligned}$$

From Lemma 3.9 Balan (2016), note that $tr\{(\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*)^2\} = [\|\mathbf{x}\|_2^4 + \|\mathbf{y}\|_2^4 - |\langle \mathbf{x}, \mathbf{y} \rangle|^2]$, where $tr\{\cdot\}$ represents the trace of a matrix. Since $\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*$ is a Hermitian matrix $tr\{(\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*)^2\} = \|\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\|_F^2$. Hence, finally we can state,

$$\mathbb{E}[Y] = \|\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\|_F^2$$

Next we focus on obtaining concentration bounds. Just as with expectation $\mathbb{E}[Y]$, we evaluate the behaviour of deviation in each individual set B, C, D and E.

1. Set B,

$$\begin{aligned}
&\sum_{(i,j,k,l) \in B} a_{ij} a_{kl} (x_i \bar{x}_j - y_i \bar{y}_j)(x_k \bar{x}_l - y_k \bar{y}_l) \\
&- \mathbb{E} \left(\sum_{(i,j,k,l) \in B} a_{ij} a_{kl} (x_i \bar{x}_j - y_i \bar{y}_j)(x_k \bar{x}_l - y_k \bar{y}_l) \right) \\
&= \sum_{i=1}^{n} \left(|a_{ii}|^2 - \sigma^2 \right) \left(|x_i|^4 + |y_i|^4 - 2|x_i|^2 |y_i|^2 \right)
\end{aligned}$$

   Note that $\forall i \in [1, n], |a_{ii}|^2 - \sigma^2$ is a centered subexponential random variable.

2. Set C,

$$\begin{aligned}
&\sum_{(i,j,k,l) \in C} a_{ij} a_{kl} (x_i \bar{x}_j - y_i \bar{y}_j)(x_k \bar{x}_l - y_k \bar{y}_l) \\
&- \mathbb{E} \left(\sum_{(i,j,k,l) \in C} a_{ij} a_{kl} (x_i \bar{x}_j - y_i \bar{y}_j)(x_k \bar{x}_l - y_k \bar{y}_l) \right) \\
&= \sum_{i,j=1, i \neq j}^{n} (|a_{ij}|^2 - \sigma^2) \\
&(|x_i|^2 |x_j|^2 - y_i \bar{y}_j x_j \bar{x}_i - x_i \bar{x}_j y_j \bar{y}_i + |y_i|^2 |y_j|^2)
\end{aligned}$$

   Again, note that $\forall i, j \in [1, n]^2, i \neq j, |a_{ij}|^2 - \sigma^2$ is a centered subexponential random variable.

3. Set D,

$$\sum_{(i,j,k,l)\in D} a_{ij}a_{kl}(x_i\bar{x}_j - y_i\bar{y}_j)(x_k\bar{x}_l - y_k\bar{y}_l)$$
$$- \mathbb{E}\left(\sum_{(i,j,k,l)\in D} a_{ij}a_{kl}(x_i\bar{x}_j - y_i\bar{y}_j)(x_k\bar{x}_l - y_k\bar{y}_l)\right)$$
$$= \sum_{i,j=1,i\neq j}^{n} (a_{ij})^2(x_i)^2(\bar{x}_j)^2$$

Note that $a_{ij}^2 = (a_{ij}^r)^2 - (a_{ij}^i)^2 + ia_{ij}^r a_{ij}^i$. This makes it easier to argue that $\forall i,j \in [1,n]^2, i \neq j, (a_{ij})^2$ is a centered subexponential random variable.

4. For elements in set E,

$$\sum_{(i,j,k,l)\in E} a_{ij}a_{kl}(x_i\bar{x}_j - y_i\bar{y}_j)(x_k\bar{x}_l - y_k\bar{y}_l)$$
$$- \mathbb{E}\left(\sum_{(i,j,k,l)\in E} a_{ij}a_{kl}(x_i\bar{x}_j - y_i\bar{y}_j)(x_k\bar{x}_l - y_k\bar{y}_l)\right)$$
$$= \sum_{(i,j,k,l)\in E} a_{ij}a_{kl}(x_i\bar{x}_j - y_i\bar{y}_j)(x_k\bar{x}_l - y_k\bar{y}_l)$$

Since a_{ij}, a_{kl} for $(i,j,k,l) \in E$ are independent, it can be easily seen that $a_{ij}a_{kl}$ is a centered subexponential random variable $\forall (i,j,k,l) \in E$.

Take $\sigma^2 = 1$. We then have the Bernstein type inequality Vershynin (2010) as,

$$\mathbb{P}\left(\left|\sum_{d=1}^{n}\frac{1}{m}|\langle A_d, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2 - \|\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\|_F^2\right| \leq t\right)$$
$$\geq 1 - c_0\exp\left(-c_1 m \min\left\{\frac{t^2}{K_4^2\|\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\|_2^4}, \frac{t}{K_4\|\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\|_\infty^2}\right\}\right)$$

for some constants $c_0, c_1 > 0$.

We introduce the normalized variable $\epsilon = \frac{t}{\|\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\|_F^2}$,

$$\mathbb{P}\left(\left|\sum_{d=1}^{n}\frac{1}{m}|\langle A_d, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2 - \|\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\|_F^2\right|\right.$$
$$\left.\geq \epsilon\|\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\|_F^2\right) \leq c_0\exp^{-c_1 m E(\epsilon)}$$

where $E(\epsilon) := \min\left\{\frac{\epsilon^2}{K^2}, \frac{\epsilon}{K}\right\}$.

Note that $\|\mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\|_F^2$ is the distance metric $d(\cdot,\cdot)$ defined in (3.2). Hence we can rewrite the high probability result more consicely as,

$$\begin{aligned}&\mathbb{P}\left(\left|\sum_{d=1}^{n}\frac{1}{m}|\langle A_d, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2 - d(\mathbf{x},\mathbf{y})^2\right|\right.\\&\left.\geq \epsilon d(\mathbf{x},\mathbf{y})^2\right) \leq c_0 \exp^{-c_1 m E(\epsilon)}\end{aligned}$$

□

Lemma B.4.2. *Given $\delta > 0$, let $\mathcal{N}_\delta$ be the smallest collection of n-dimensional balls of radius δ whose union covers the sphere S^{n-1}. Then, for any matrix $A \in \mathbb{C}^{n\times n}$, we have*

$$\begin{aligned}&(1-2\delta)\sup_{\mathbf{x}_1,\mathbf{x}_2\in S^{n-1}}|\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle|\\&\leq \sup_{\mathbf{x}_1,\mathbf{x}_2\in \mathcal{N}_\delta}|\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle|\\&\leq (1+2\delta)\sup_{\mathbf{x}_1,\mathbf{x}_2\in S^{n-1}}|\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle|.\end{aligned}$$

Proof. In the proof, we relate the supremum of $|\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle|$ over $\mathbf{x},\mathbf{y} \in S^{n-1}$ to its supremum over $\mathbf{x},\mathbf{y} \in \mathcal{N}_\delta$.

Since $\mathcal{N}_\delta$ covers S^{n-1}, $\forall \mathbf{x} \in S^{n-1}$, $\exists \mathbf{u} \in \mathcal{N}_\delta$ such that $\|\mathbf{x}-\mathbf{u}\| \leq \delta$.

Hence $\forall \mathbf{x}_1, \mathbf{x}_2 \in S^{n-1}$, $\exists \mathbf{y}_1, \mathbf{y}_2 \in \mathcal{N}_\delta$ such that,

$$\begin{aligned}&|\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle - \langle A, \mathbf{y}_1\mathbf{y}_1^* - \mathbf{y}_2\mathbf{y}_2^*\rangle|\\&= |\langle A\mathbf{x}_1, \mathbf{x}_1\rangle - \langle A\mathbf{y}_1, \mathbf{y}_1\rangle| - \langle A\mathbf{x}_2, \mathbf{x}_2\rangle - \langle A\mathbf{y}_2, \mathbf{y}_2\rangle|\\&= |\langle A\mathbf{x}_1, \mathbf{x}_1\rangle - \langle A\mathbf{x}_1, \mathbf{y}_1\rangle + \langle A\mathbf{x}_1, \mathbf{y}_1\rangle - \langle A\mathbf{y}_1, \mathbf{y}_1\rangle|\\&+ |\langle A\mathbf{x}_2, \mathbf{x}_2\rangle - \langle A\mathbf{x}_2, \mathbf{y}_2\rangle + \langle A\mathbf{x}_2, \mathbf{y}_2\rangle - \langle A\mathbf{y}_2, \mathbf{y}_2\rangle|\\&= |\langle A\mathbf{x}_1, \mathbf{x}_1 - \mathbf{y}_1\rangle + \langle A\mathbf{y}_1, \mathbf{x}_1 - \mathbf{y}_1\rangle|\\&+ |\langle A\mathbf{x}_2, \mathbf{x}_2 - \mathbf{y}_2\rangle + \langle A\mathbf{y}_2, \mathbf{x}_2 - \mathbf{y}_2\rangle|\\&\leq 2\|A\|\|\mathbf{x}_1 - \mathbf{y}_1\| + 2\|A\|\|\mathbf{x}_2 - \mathbf{y}_2\|\\&\leq 4\delta\|A\|\end{aligned}$$

where $\|A\|$ denotes the spectral norm of the matrix A, i.e.,

$$\begin{aligned}\|A\| &= \sup_{\mathbf{x}\in S^{n-1}}|\langle A\mathbf{x}, \mathbf{x}\rangle|\\&= \frac{1}{2}\sup_{\mathbf{x}\in S^{n-1}}|\langle A\mathbf{x}, \mathbf{x}\rangle| + \frac{1}{2}\sup_{\mathbf{y}\in S^{n-1}}|\langle A\mathbf{y}, \mathbf{y}\rangle|\\&= \frac{1}{2}\sup_{\mathbf{x},\mathbf{y}\in S^{n-1}}|\langle A, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|\end{aligned}$$

We conclude,

$$\begin{aligned}&|\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle| - |\langle A, \mathbf{y}_1\mathbf{y}_1^* - \mathbf{y}_2\mathbf{y}_2^*\rangle| \leq 4\delta\|A\|\\&|\langle A, \mathbf{y}_1\mathbf{y}_1^* - \mathbf{y}_2\mathbf{y}_2^*\rangle| \geq |\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle| - 4\delta\|A\|\end{aligned}$$

And,

$$|\langle A, \mathbf{y}_1\mathbf{y}_1^* - \mathbf{y}_2\mathbf{y}_2^*\rangle| - |\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle| \leq 4\delta\|A\|$$
$$|\langle A, \mathbf{y}_1\mathbf{y}_1^* - \mathbf{y}_2\mathbf{y}_2^*\rangle| \leq |\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle| + 4\delta\|A\|$$

Taking supremum,

$$\begin{aligned}
&\sup_{\mathbf{x}\in\mathcal{N}_\delta} |\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle| \\
&\geq \sup_{\mathbf{x}\in S^{n-1}} |\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle| - 4\delta\|A\| \\
&= (2 - 4\delta)\|A\| \\
&= (1 - 2\delta) \sup_{\mathbf{x}\in S^{n-1}} |\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle| \\
&\sup_{\mathbf{x}\in\mathcal{N}_\delta} |\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle| \\
&\leq \sup_{\mathbf{x}\in S^{n-1}} |\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle| + 4\delta\|A\| \\
&= (2 + 4\delta)\|A\| \\
&= (1 + 2\delta) \sup_{\mathbf{x}\in S^{n-1}} |\langle A, \mathbf{x}_1\mathbf{x}_1^* - \mathbf{x}_2\mathbf{x}_2^*\rangle|
\end{aligned}$$

□

Theorem B.4.3. *Let $\mathcal{A} = \{A_i\}_{i=1}^m$ be the set of complex Gaussian random matrices, and assume the number of measurements satisfies $m > Cn$. Then, for any given $\xi \in (0,1)$, there exist constants $C, c_0, d_0 > 0$ and $\beta \geq \alpha > 0$ such that, with probability at least $1 - \xi$, the following relation holds*

$$\alpha d(\mathbf{x}, \mathbf{y}) \leq \|\mathcal{M}_\mathcal{A}(\mathbf{x}) - \mathcal{M}_\mathcal{A}(\mathbf{y})\|_2 \leq \beta d(\mathbf{x}, \mathbf{y}).$$

Proof. Consider $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$. If $\mathbf{x} \sim \mathbf{y}$, then $d(\mathbf{x}, \mathbf{y}) = 0$ and $\mathcal{M}_\mathcal{A}(\mathbf{x}) = \mathcal{M}_\mathcal{A}(\mathbf{y})$, and the result holds trivially. Therefore, in the sequel, we assume that the vectors are distinct, implying that $d(\mathbf{x}, \mathbf{y}) > 0$.

From Lemma B.4.1, for a given $\epsilon > 0$ we have

$$\mathbb{P}\left(\left|\sum_{i=1}^m \frac{1}{m} \frac{|\langle A_i, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2}{d(\mathbf{x}, \mathbf{y})^2} - 1\right| \geq \epsilon\right) \leq de^{-cm\epsilon}.$$

According to Vershynin (2010), for any $\delta > 0$, we have the following upper bound on the size of the covering: $|\mathcal{N}_\delta| \leq \left(\frac{12}{\delta}\right)^n$. Therefore, for a given $\epsilon, \delta > 0$, by Lemma B.4.1 and the preceding union bound, we have that

$$\begin{aligned}
&\mathbb{P}\left(\sup_{\mathbf{x},\mathbf{y}\in\mathcal{N}_\delta} \left|\sum_{i=1}^m \frac{1}{m} \frac{|\langle A_i, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2}{d(\mathbf{x}, \mathbf{y})^2} - 1\right| > \epsilon\right) \\
&\leq de^{-cm\epsilon} \left(\frac{12}{\delta}\right)^n,
\end{aligned}$$

where d, c are the same constants as in Lemma B.4.1. This implies that

$$\mathbb{P}\left(\sup_{\mathbf{x},\mathbf{y}\in\mathcal{N}_\delta}\left|\sum_{i=1}^m \frac{1}{m}\frac{|\langle A_i, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2}{d(\mathbf{x},\mathbf{y})^2} - 1\right| \leq \epsilon\right)$$
$$\geq 1 - de^{-cm\epsilon}\left(\frac{12}{\delta}\right)^n.$$

Now, observe that

$$\sup_{\mathbf{x},\mathbf{y}\in\mathcal{N}_\delta}\left|\sum_{i=1}^m \frac{1}{m}\frac{|\langle A_i, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2}{d(\mathbf{x},\mathbf{y})^2} - 1\right|$$
$$\geq \sup_{\mathbf{x},\mathbf{y}\in\mathcal{N}_\delta}\sum_{i=1}^m \frac{1}{m}\frac{|\langle A_i, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2}{d(\mathbf{x},\mathbf{y})^2} - 1.$$

Therefore,

$$\mathbb{P}\left(\sup_{\mathbf{x},\mathbf{y}\in\mathcal{N}_\delta}\sum_{i=1}^m \frac{1}{m}\frac{|\langle A_i, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2}{d(\mathbf{x},\mathbf{y})^2} - 1 \leq \epsilon\right)$$
$$\geq 1 - de^{-cm\epsilon}\left(\frac{12}{\delta}\right)^n.$$

By applying the covering result of Lemma B.4.2 to each matrix A_i, averaging the resulting relations over m, and using

$$\sup_{\mathbf{x},\mathbf{y}\in\mathcal{N}_\delta} d(\mathbf{x},\mathbf{y}) \leq (1+2\delta)\sup_{\mathbf{x},\mathbf{y}\in S^{n-1}} d(\mathbf{x},\mathbf{y}).$$

we obtain

$$\sup_{\mathbf{x},\mathbf{y}\in\mathcal{N}_\delta}\sum_{i=1}^m \frac{1}{m}\frac{|\langle A_i, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2}{d(\mathbf{x},\mathbf{y})^2}$$
$$\geq \sup_{\mathbf{x},\mathbf{y}\in S^{n-1}}\sum_{i=1}^m \frac{(1-2\delta)^2}{(1+2\delta)^2 m}\frac{|\langle A_i, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2}{d(\mathbf{x},\mathbf{y})^2}.$$

Thus, we can conclude that

$$\mathbb{P}\left(\sup_{\mathbf{x},\mathbf{y}\in S^{n-1}}\sum_{i=1}^m \frac{1}{m}\frac{|\langle A_i, \mathbf{x}\mathbf{x}^* - \mathbf{y}\mathbf{y}^*\rangle|^2}{d(\mathbf{x},\mathbf{y})^2} \leq \frac{(1+2\delta)^2(1+\epsilon)}{(1-2\delta)^2}\right)$$
$$\geq 1 - de^{-cm\epsilon}\left(\frac{12}{\delta}\right)^n.$$

Similarly, we can prove that

$$\mathbb{P}\left(\inf_{\mathbf{x},\mathbf{y}\in S^{n-1}}\sum_{i=1}^{m}\frac{1}{m}\frac{|\langle A_i,\mathbf{x}\mathbf{x}^*-\mathbf{y}\mathbf{y}^*\rangle|^2}{d(\mathbf{x},\mathbf{y})^2}\leq\frac{(1-2\delta)^2(1-\epsilon)}{(1+2\delta)^2}\right)$$
$$\geq 1-de^{-cm\epsilon}\left(\frac{12}{\delta}\right)^n.$$

Letting C be such that $C>\frac{\log 12d-\log\delta\xi}{c\epsilon}$ and letting $m\geq Cn$, we can see that the following relation holds with probability at least $1-\xi$: for all $\mathbf{x},\mathbf{y}\in\mathbb{C}^n$,

$$\beta d(\mathbf{x},\mathbf{y})\geq\|\mathcal{M}_{\mathcal{A}}(\mathbf{x})-\mathcal{M}_{\mathcal{A}}(\mathbf{y})\|_2\geq\alpha d(\mathbf{x},\mathbf{y}),$$

where α,β are given by

$$\alpha\triangleq\frac{((1-2\delta)^2(1-\epsilon)}{(1+2\delta)^2},\quad\beta\triangleq\frac{((1+2\delta)^2(1+\epsilon)}{(1-2\delta)^2}.$$

Notice that we essentially have a choice of the values of δ and ϵ. The closer they are to 0, the stronger the stability result. However, this also implies the larger m, the number of measurements, needs to be. □

We prove the following result for showing κ-phase-discriminating property, $\kappa\geq 0$, is satisfied for a wide variety of random matrices (further used in Theorem 3.6.7)

Lemma B.4.4. *Let $\{A_d\}_{d=1}^n$ be a set of Hermitian Gaussian random matrices. Then,*

$$\mathbb{P}\Bigg(\Bigg|\frac{1}{m}\sum_{d=1}^{m}\left(\langle A_d,\Delta\bar{\Delta}^\top\rangle\langle A_d,\mathbf{x}\bar{\mathbf{x}}^\top\rangle-\langle A_d,\Delta\bar{\mathbf{x}}^\top\rangle\langle A_d,\mathbf{x}\bar{\Delta}^\top\rangle\right)$$
$$+\|\Delta\mathbf{x}^T-\mathbf{x}\Delta^T\|_F^2\Big|\leq t\|\Delta\mathbf{x}^T-\mathbf{x}\Delta^T\|_F^2\Bigg)$$
$$\geq 1-c_0exp\left(-c_1mD(t)\right)$$

where,

$$D(t):=\min\left\{\frac{t^2}{K_4^2},\frac{t}{K_4}\right\}$$

Proof. Let $A\in\mathbb{C}^{n\times n}$ be a complex Hermitian Gaussian random matrix, i.e.

1. $\forall\ i,\ a_{ii}\sim\mathcal{N}(0,\sigma^2)$.
2. $\forall\ i,j,\ i\neq j,\ a_{ij}\sim\mathcal{N}(0,\frac{\sigma^2}{2})+i\mathcal{N}(0,\frac{\sigma^2}{2})$.

Define the random variable Y,

$$\begin{aligned}Y&=\frac{1}{m}\sum_{d=1}^{m}\langle A_d,\Delta\bar{\Delta}^\top\rangle\langle A_d,\mathbf{x}\bar{\mathbf{x}}^\top\rangle-\langle A_d,\Delta\bar{\mathbf{x}}^\top\rangle\langle A_d,\mathbf{x}\bar{\Delta}^\top\rangle\\&=\frac{1}{m}\sum_{d=1}^{m}\left(\sum_{ij}a_{ij}^d\Delta_i\bar{\Delta}_j\right)\left(\sum_{kl}a_{kl}^d\mathbf{x}_k\bar{\mathbf{x}}_l\right)\\&-\left(\sum_{ij}a_{ij}^d\mathbf{x}_i\bar{\Delta}_j\right)\left(\sum_{kl}a_{kl}^d\Delta_k\bar{\mathbf{x}}_l\right)\end{aligned}$$

For any A_d, Split the entire summation $(i,j,k,l) \in [1,n]^4$ into the following 4 sets such that:

1. $A := \{(i,j,k,l) | i=j=k=l\}$
2. $B := \{(i,j,k,l) | i=k, j=l\} \cap A^C$
3. $C := \{(i,j,k,l) | i=l, j=k\} \cap A^C$
4. $D := \{(i,j,k,l)\} \cap A^C \cap B^C \cap C^C$

Calculating the expectation of the sum of the elements in each individual sets:

1. For set A,

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{ij} a_{ij}\Delta_i\bar{\Delta}_j\right)\left(\sum_{kl} a_{kl}\mathbf{x}_k\bar{\mathbf{x}}_l\right)\right. \\
&\left.-\left(\sum_{ij} a_{ij}\mathbf{x}_i\bar{\Delta}_j\right)\left(\sum_{kl} a_{kl}\Delta_k\bar{\mathbf{x}}_l\right)\right] \\
&= \mathbb{E}\left[a_{ii}^2\Delta_i\bar{\Delta}_i\mathbf{x}_i\bar{\mathbf{x}}_j - a_{ii}^2\Delta_i\bar{\Delta}_i\mathbf{x}_i\bar{\mathbf{x}}_i\right] = \mathbb{E}[0] = 0
\end{aligned}$$

2. For set B,

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{ij} a_{ij}\Delta_i\bar{\Delta}_j\right)\left(\sum_{kl} a_{kl}\mathbf{x}_k\bar{\mathbf{x}}_l\right)\right. \\
&\left.-\left(\sum_{ij} a_{ij}\mathbf{x}_i\bar{\Delta}_j\right)\left(\sum_{kl} a_{kl}\Delta_k\bar{\mathbf{x}}_l\right)\right] \\
&= \mathbb{E}\left[a_{ij}^2\Delta_i\bar{\Delta}_j\mathbf{x}_i\bar{\mathbf{x}}_j - a_{ij}^2\Delta_i\bar{\Delta}_j\mathbf{x}_i\bar{\mathbf{x}}_j\right] = \mathbb{E}[0] = 0
\end{aligned}$$

3. For set C, since the matrix A is hermitian $a_{ij} = \bar{a}_{ji}$

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{ij} a_{ij}\Delta_i\bar{\Delta}_j\right)\left(\sum_{kl} a_{kl}\mathbf{x}_k\bar{\mathbf{x}}_l\right)\right. \\
&\left.-\left(\sum_{ij} a_{ij}\mathbf{x}_i\bar{\Delta}_j\right)\left(\sum_{kl} a_{kl}\Delta_k\bar{\mathbf{x}}_l\right)\right] \\
&= \left[|a_{ij}|^2\Delta_i\bar{\Delta}_j\mathbf{x}_j\bar{\mathbf{x}}_i - |a_{ij}|^2\Delta_j\bar{\Delta}_j\mathbf{x}_i\bar{\mathbf{x}}_i\right]
\end{aligned}$$

   Notice that $\forall i,j$

$$\begin{aligned}
&= |a_{ji}|^2\Delta_j\bar{\Delta}_i\mathbf{x}_i\bar{\mathbf{x}}_j + |a_{ij}|^2\Delta_i\bar{\Delta}_j\mathbf{x}_j\bar{\mathbf{x}}_i \\
&- |a_{ji}|^2\Delta_i\bar{\Delta}_i\mathbf{x}_j\bar{\mathbf{x}}_j - |a_{ij}|^2\Delta_j\bar{\Delta}_j\mathbf{x}_i\bar{\mathbf{x}}_i
\end{aligned}$$

Since $|a_{ij}|^2 = |a_{ji}|^2$. Thus,

$$\begin{aligned}
&= |a_{ji}|^2 \ \left[\Delta_j \bar{\Delta}_i \mathbf{x}_i \bar{\mathbf{x}}_j + \Delta_i \bar{\Delta}_j \mathbf{x}_j \bar{\mathbf{x}}_i \right. \\
&\qquad \left. -\Delta_i \bar{\Delta}_i \mathbf{x}_j \bar{\mathbf{x}}_j - \Delta_j \bar{\Delta}_j \mathbf{x}_i \bar{\mathbf{x}}_i \right] \\
&= |a_{ji}|^2 \ \left[\Delta_j \bar{\Delta}_i \mathbf{x}_i \bar{\mathbf{x}}_j + \Delta_i \bar{\Delta}_j \mathbf{x}_j \bar{\mathbf{x}}_i \right. \\
&\qquad \left. -|\Delta_i|^2 |\mathbf{x}_j|^2 - |\Delta_j|^2 |\mathbf{x}_i|^2 \right] \\
&= |a_{ji}|^2 \ \left[\Delta_j \bar{\Delta}_i \mathbf{x}_i \bar{\mathbf{x}}_j + \Delta_i \bar{\Delta}_j \mathbf{x}_j \bar{\mathbf{x}}_i \right. \\
&\qquad \left. -|\Delta_i|^2 |\mathbf{x}_j|^2 - |\Delta_j|^2 |\mathbf{x}_i|^2 \right] \\
&= -|a_{ij}|^2 | \ \Delta_i \mathbf{x}_j - \mathbf{x}_i \Delta_j|^2
\end{aligned}$$

Thus,

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{ij} a_{ij} \Delta_i \bar{\Delta}_j\right)\left(\sum_{kl} a_{kl} \mathbf{x}_k \bar{\mathbf{x}}_l\right)\right. \\
&\left. - \left(\sum_{ij} a_{ij} \mathbf{x}_i \bar{\Delta}_j\right)\left(\sum_{kl} a_{kl} \Delta_k \bar{\mathbf{x}}_l\right)\right] \\
&= -\sigma^2 \|\Delta \mathbf{x}^T - \mathbf{x} \Delta^T\|_F^2
\end{aligned} \tag{B.14}$$

4. For set D, as all the elements $(i, j, k, l) \in D$ make a_{ij}, a_{kl} independent of each other, we have,

$$\mathbb{E}\left(\sum a_{ij} a_{kl} \bar{\Delta}_j \bar{\mathbf{x}}_l \left(\Delta_i \mathbf{x}_k - \Delta_k \mathbf{x}_i\right)\right) = 0$$

Hence we can conclude,

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{m} \sum_{d=1}^{m} \left(\langle A_d, \Delta \bar{\Delta}^\top \rangle \langle A_d, \mathbf{x} \bar{\mathbf{x}}^\top \rangle - \langle A_d, \Delta \bar{\mathbf{x}}^\top \rangle \langle A_d, \mathbf{x} \bar{\Delta}^\top \rangle\right)\right] \\
&= -\sigma^2 \|\Delta \mathbf{x}^T - \mathbf{x} \Delta^T\|_F^2
\end{aligned}$$

We focus our attention on obtaining concentration bounds. Evaluating the behaviour on the elements in set D,

$$\begin{aligned}
&\frac{1}{m} \sum_{d=1}^{m} \left(\sum_{(i,j,k,l) \in D} a_{ij}^d a_{kl}^d \bar{\Delta}_j \bar{\mathbf{x}}_l \left(\Delta_i \mathbf{x}_k - \Delta_k \mathbf{x}_i\right) \right) \\
&- \mathbb{E}\left(\frac{1}{m} \sum_{d=1}^{m} \sum_{(i,j,k,l) \in D} a_{ij}^d a_{kl}^d \bar{\Delta}_j \bar{\mathbf{x}}_l \left(\Delta_i \mathbf{x}_k - \Delta_k \mathbf{x}_i\right) \right) \\
&= \sum_{(i,j,k,l) \in D} a_{ij}^d a_{kl}^d \bar{\Delta}_j \bar{\mathbf{x}}_l \left(\Delta_i \mathbf{x}_k - \Delta_k \mathbf{x}_i\right)
\end{aligned}$$

We can see that the above is a centered subexponential random variable.

Take $\sigma^2 = 1$. We then have the Bernstein type inequality Vershynin (2010) as,

$$\begin{aligned}
&\mathbb{P}\left(\left|\frac{1}{m}\sum_{d=1}^{m}\left(\langle A_d, \Delta\bar{\Delta}^\top\rangle\langle A_d, \mathbf{x}\bar{\mathbf{x}}^\top\rangle - \langle A_d, \Delta\bar{\mathbf{x}}^\top\rangle\langle A_d, \mathbf{x}\bar{\Delta}^\top\rangle\right)\right.\right.\\
&\left.\left. + \ \|\Delta\mathbf{x}^T - \mathbf{x}\Delta^T\|_F^2\right| \le t\|\Delta\mathbf{x}^T - \mathbf{x}\Delta^T\|_F^2\right)\\
&\ge 1 - c_0\exp\left(-c_1 m\min\left\{\frac{t^2}{K_4^2}, \frac{t}{K_4}\right\}\right)
\end{aligned}$$

□

B.5 Appendix : Non-convex Landscape for Gaussian Random $\mathcal{A}$

B.5.1 Proof of Theorem 3.6.7

Theorem B.5.1. *Let $\{A_i\}_{i=1}^m$ be a set of complex $n \times n$ Gaussian random matrices, and let $m > Cn$ for some constant $C > 0$. Let the scalars $\{c_i\}_{i=1}^m$ characterizing the objective function f of problem (P2) be generated by quadratic measurements of an unknown vector $\mathbf{z}$. Then, for any given $\xi \in (0,1)$, there exist positive constants β, γ, and ζ such that the following statements hold with probability at least $1 - \xi$:*

1) The function f is (β, ζ, γ)-strict saddle, and

2) Every local minimum $\mathbf{w}$ of f satisfies $d(\mathbf{w}, \mathbf{z}) = 0$

Proof. Notice that,

$$\begin{aligned}
&(\Delta A\bar{\mathbf{x}})^2 + (\bar{\Delta}^\top A\mathbf{x})^2\\
&= \left(\langle A, \mathbf{x}\bar{\Delta}^\top + \Delta\bar{\mathbf{x}}^\top\rangle\right)^2 - 2(\Delta^\top A\bar{\mathbf{x}})(\bar{\Delta}^\top A\mathbf{x})\\
&= \left(\langle A, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top + \Delta\bar{\Delta}^\top\rangle\right)^2 - 2(\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)
\end{aligned}$$

Using (B.11), we can reorganize,

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_i(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\\
&= \langle A_i, 2\mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, 2\Delta\bar{\Delta}^\top\rangle - 2(\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)\\
&+ \left(\langle A, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top + \Delta\bar{\Delta}^\top\rangle\right)^2\\
&= 2\left(\langle A_i, 2\mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle\right)\\
&+ \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle + \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle\\
&+ \langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top + \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\\
&- 2(\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)
\end{aligned}$$

Adding and subtracting $2\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle$, reorganizing,

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_i(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\\
&= 2\langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top\rangle\\
&+ 2\left(\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top + \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\right)\\
&- 2\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\\
&+ \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\\
&+ \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle - 2(\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)\\
&+ \langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top + \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle
\end{aligned}$$

Adding and subtracting $\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle$, reorganizing,

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_i(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\\
&= 2\langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top\rangle - 2(\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)\\
&- 3\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\\
&+ \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle\\
&+ 4\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top + \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle
\end{aligned}$$

Using equation (B.12),

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_i(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\\
&= 2\langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top\rangle - 2(\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)\\
&- 3\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\\
&+ \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle + 4\langle \nabla g_i(\mathbf{x}, \bar{\mathbf{x}}), \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\rangle
\end{aligned}$$

Overall, we can conclude that,

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix} = \frac{1}{m}\sum_{i=1}^m \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_i(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\\
&= \frac{2}{m}\sum_{i=1}^m \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top\rangle - (\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle)\\
&- \frac{3}{m}\sum_{i=1}^m \langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\\
&+ \frac{1}{m}\sum_{i=1}^m \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle + \frac{4}{m}\sum_{i=1}^m \langle \nabla g_i(\mathbf{x}, \bar{\mathbf{x}}), \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\rangle
\end{aligned}$$

Using Lemma B.4.1 and Lemma B.4.4, we can conclude that with probability greater than $1 - c_1 e^{-c_2 m \min\{D(t), E(\epsilon)\}}$,

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix} \\
&\leq -(1-t)\|\Delta\mathbf{x}^T - \mathbf{x}\Delta^T\|_F^2 + 4\delta\|\Delta\|_2 + \beta\|\Delta\Delta^*\|_F^2 \\
&- 3\alpha\|\mathbf{x}\mathbf{x}^* - \mathbf{z}(\mathbf{z})^*\|_F^2 \\
&\leq -(1-t)\|\Delta\mathbf{x}^T - \mathbf{x}\Delta^T\|_F^2 + 4\delta\|\Delta\|_2 \\
&+ 2\beta\|\mathbf{x}\mathbf{x}^* - \mathbf{z}(\mathbf{z})^*\|_F^2 - 3\alpha\|\mathbf{x}\mathbf{x}^* - \mathbf{z}(\mathbf{z})^*\|_F^2 \\
&\leq (2\beta - 3\alpha)\|\mathbf{x}\mathbf{x}^* - \mathbf{z}(\mathbf{z})^*\|_F^2 + 4\delta\|\Delta\|_2 \\
&- (1-t)\|\Delta\mathbf{x}^T - \mathbf{x}\Delta^T\|_F^2
\end{aligned} \tag{B.15}$$

where $\exists c_1, c_2 > 0$ which can be computed from Lemma B.4.1 and Lemma B.4.4.

For any $\xi \in [0, 1]$, there can be multiple possibilities of the constants β, ζ and γ satisfying Theorem 3.6.7.

For instance : Given ξ, we can take enough measurements $m = O(n)$ such that the mapping $\mathcal{M}_{\mathcal{A}}$ is $(1-c, 1+c)$-stable, for some small $c > 0$ and $t \leq 1$. Suppose the current vector $\mathbf{x}$ is not close to $\mathbf{x}^*$ such that $\|\Delta\| \geq C_0\delta$, for sufficiently large $C_0 > 0$, then we have

$$\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix} \leq (-1+5c)C_0^2\delta^2 + 4C_0\delta^2 \leq 0$$

A particular set of (c, C_0) which fit the above condition is $c = \frac{1}{20}, C_0 = 10$, then

$$\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix} \leq (-1+5c)C_0^2\delta^2 + 4C_0\delta^2 \leq 0$$

Hence we can conclude that the function $f(\mathbf{x}) = g(\mathbf{x}, \overline{\mathbf{x}})$ satisfies at-least one of the following is true,

- $\|\nabla g(\mathbf{x})\| \geq \delta$
- For the direction vector Δ,

$$\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix} \leq (-1+5c)C_0^2\delta^2 + 4C_0\delta^2$$

- $d(\mathbf{x}, \mathbf{z}) \leq C_0\delta$

Thus there exists constants $\beta, \zeta, \gamma > 0$ such that the function $f(\mathbf{x})$ is (β, ζ, γ) strict saddle.

Following up on equation (B.28), the only possible way that the hessain $\nabla^2 g(\mathbf{x}, \overline{\mathbf{x}}) \succeq 0$ is if $\|\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^*\| = 0$. Hence we can conclude that all local minimas, i.e. all $\mathbf{w}$ such that $\nabla^2 g(\mathbf{w}, \overline{\mathbf{w}}) \succeq 0$ has to satisfy $\|\mathbf{w}\mathbf{w}^* - \mathbf{z}\mathbf{z}^*\| = 0$ and hence satisfies $\mathbf{z} \sim \mathbf{w}$ which makes $\mathbf{w}$ the solution of the problem (P1).

□

B.6 Appendix : Robustness Proofs

The following lemma will be used in later used to prove Theorem 3.7.1,

Lemma B.6.1. *For any Hermitian matrix M, Gaussian noise η_i and complex Hermitian Gaussian matrices $\{A_i\}_{i=1}^m$, we can say that*

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m}\eta_i\langle A_i, M\rangle\right| \geq \epsilon_\eta\sigma_\eta\|M\|_F\right) \leq 2e^{\left(-cm\min\left\{\frac{\epsilon_\eta^2}{K_5^2}, \frac{\epsilon_\eta}{K_5}\right\}\right)} \tag{B.16}$$

for some constants $c, K_5 > 0$.

Proof. For all $k \in [m]$, let $A_k \in \mathbb{C}^{n\times n}$ be complex Hermitian Gaussian random matrix, i.e.

1. $\forall\ i,\ a_{ii} \sim \mathcal{N}(0, \sigma^2)$.
2. $\forall\ i, j,\ i \neq j,\ a_{ij} \sim \mathcal{N}(0, \frac{\sigma^2}{2}) + i\mathcal{N}(0, \frac{\sigma^2}{2})$.

and $\eta_k \sim \mathcal{N}(0, \sigma_\eta^2)$ be Gaussian random variables. Define the random variable Y,

$$Y = \left|\frac{1}{m}\sum_{k=1}^{m}\eta_k\langle A_k, X\rangle\right| \tag{B.17}$$

We first quantify the nature of $\eta_k\langle A_k, M\rangle$

$$\begin{aligned}
\eta_k\langle A_k, M\rangle &= \eta_k\sum_i a_{ii}m_{ii} + \eta_k\sum_{i,j,i\neq j} a_{ij}m_{ij} \\
&= \eta_k\sum_i a_{ii}m_{ii} + \eta_k\sum_{i,j,i>j} a_{ij}m_{ij} + \eta_k\sum_{i,j,i>j}\overline{a_{ij}m_{ij}} \\
&= \sum_i \eta_k a_{ii}m_{ii} + 2\sum_{i,j,i\neq j}\eta_k Re(a_{ij}m_{ij})
\end{aligned}$$

Hence we can say that,

$$\mathbb{E}[\eta_k\langle A_k, M\rangle] = 0 \tag{B.18}$$

and,

$$\mathbb{E}[(\eta_k\langle A_k, M\rangle)^2] = \sigma_\eta^2\|M\|_F^2 \tag{B.19}$$

Note that $\langle A_k, M\rangle \sim \mathcal{N}(0, \|M\|_F^2)$ and $\eta_k \sim \mathcal{N}(0, \sigma_\eta^2)$ and hence we can conclude that $\eta_i\langle A_k, M\rangle$ is a centered subexponential random variable. Applying Bernstein inequality, we have,

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m}\eta_i\langle A_i, M\rangle\right| \geq t\right) \leq 2e^{\left(-cm\min\left\{\frac{t^2}{K_5^2\sigma_\eta^2\|M\|_2^2}, \frac{t}{K_5\sigma_\eta\|M\|_\infty}\right\}\right)} \tag{B.20}$$

using a substitution $\epsilon_\eta = \frac{t}{\sigma_\eta \|M\|_F}$, we have,

$$\mathbb{P}\left(\left|\frac{1}{m}\sum_{i=1}^{m}\eta_i\langle A_i, M\rangle\right| \geq \epsilon\sigma_\eta\|M\|_F\right) \leq 2e^{\left(-cm\min\left\{\frac{\epsilon_\eta^2}{K_5^2}, \frac{\epsilon_\eta}{K_5}\right\}\right)} \tag{B.21}$$

□

Theorem B.6.2. *Let $\{A_i\}_{i=1}^m$ be a set of complex $n \times n$ Gaussian random matrices, and let $m > Cn$ for some large constant $C > 0$. Let the scalars $\{c_i\}_{i=1}^m$ characterizing the objective function f of problem P1 be generated by quadratic measurements of an unknown vector $\mathbf{z}$. Let the Let $\hat{\mathbf{x}} \in \mathbb{C}^n$ be such that $\|\nabla g_\eta(\hat{\mathbf{x}})\| \leq \delta$, where g_η is the ℓ_2-loss function (F_η). Then with probability $1 - c_1e^{-c_2m\epsilon} - 2e^{-c_3m\epsilon_\eta}$ the following holds:*

$$\|\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^*\|_F \leq \frac{4(\delta + \epsilon_\eta\sigma_\eta)}{1 - 5\epsilon} \tag{B.22}$$

where $\sigma_\eta \in \mathbb{R}$ is the noise variance.

Proof. Following steps similar to that of Theorem 3.6.7, for the noisy case, we have that,

$$\begin{aligned} g_\eta(\mathbf{x}, \overline{\mathbf{x}}) &= \frac{1}{m}\sum_{i=1}^{m}\left(\langle A_i\mathbf{x}, \mathbf{x}\rangle - c_i + \eta_i\right)^2 \\ &= g_0(\mathbf{x}, \overline{\mathbf{x}}) + \underbrace{\frac{1}{m}\sum_{i=1}^{m} 2\eta_i(\langle A_i\mathbf{x}, \mathbf{x}\rangle - c_i) + \frac{1}{m}\sum_{i=1}^{m}\eta_i^2}_{N(\mathbf{x}, \bar{\mathbf{x}})} \end{aligned} \tag{B.23}$$

$$\begin{aligned} \nabla N(\mathbf{x}, \bar{\mathbf{x}}) &= \frac{1}{m}\sum_{i=1}^{m}\begin{bmatrix}\eta_i A_i\mathbf{x} \\ \eta_i A_i\bar{\mathbf{x}}\end{bmatrix} \\ \left\langle \nabla N(\mathbf{x}, \bar{\mathbf{x}}), \begin{bmatrix}\Delta \\ \bar{\Delta}\end{bmatrix}\right\rangle &= \frac{1}{m}\sum_{i=1}^{m}\eta_i\langle A_i, \Delta\bar{\mathbf{x}}^T + \mathbf{x}\bar{\Delta}^T\rangle \end{aligned} \tag{B.24}$$

Thus we have,

$$\begin{aligned} &\left\langle \nabla g_\eta(\mathbf{x}, \bar{\mathbf{x}}), \begin{bmatrix}\Delta \\ \bar{\Delta}\end{bmatrix}\right\rangle \\ &= \left\langle \nabla g_0(\mathbf{x}, \bar{\mathbf{x}}), \begin{bmatrix}\Delta \\ \bar{\Delta}\end{bmatrix}\right\rangle + \frac{1}{m}\sum_{i=1}^{m}\eta_i\langle A_i, \Delta\bar{\mathbf{x}}^T + \mathbf{x}\bar{\Delta}^T\rangle \end{aligned} \tag{B.25}$$

$$\begin{aligned} \nabla^2 N(\mathbf{x}, \bar{\mathbf{x}}) &= \frac{1}{m}\sum_{i=1}^{m}\begin{bmatrix}\eta_i A_i & 0 \\ 0 & \eta_i A_i\end{bmatrix} \\ \begin{bmatrix}\Delta \\ \bar{\Delta}\end{bmatrix}^* \nabla^2 N(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta \\ \bar{\Delta}\end{bmatrix} &= \frac{2}{m}\sum_{i=1}^{m}\eta_i\langle A_i, \Delta\bar{\Delta}\rangle \end{aligned} \tag{B.26}$$

$$
\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_\eta(\mathbf{x},\bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix} \\
&= \frac{1}{m}\sum_{i=1}^m \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_i(\mathbf{x},\bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix} + \frac{2}{m}\sum_{i=1}^m \eta_i \langle A_i, \Delta\bar{\Delta}\rangle \\
&= \frac{2}{m}\sum_{i=1}^m \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top\rangle - (\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle) \\
&- \frac{3}{m}\sum_{i=1}^m \langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle \\
&+ \frac{1}{m}\sum_{i=1}^m \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle + \frac{4}{m}\sum_{i=1}^m \langle \nabla g_i(\mathbf{x},\bar{\mathbf{x}}), \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\rangle \\
&+ \frac{2}{m}\sum_{i=1}^m \eta_i \langle A_i, \Delta\bar{\Delta}\rangle \\
&= \frac{2}{m}\sum_{i=1}^m \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top\rangle - (\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle) \\
&- \frac{3}{m}\sum_{i=1}^m \langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle \\
&+ \frac{1}{m}\sum_{i=1}^m \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle + 4\langle \nabla g_\eta(\mathbf{x},\bar{\mathbf{x}}), \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\rangle \\
&+ \frac{2}{m}\sum_{i=1}^m \eta_i \langle A_i, \Delta\bar{\Delta}\rangle - \frac{4}{m}\sum_{i=1}^m \eta_i \langle A_i, \Delta\bar{\mathbf{x}}^T + \mathbf{x}\bar{\Delta}^T\rangle \\
&= \frac{2}{m}\sum_{i=1}^m \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top\rangle - (\langle A, \Delta\bar{\mathbf{x}}^\top\rangle)(\langle A, \mathbf{x}\bar{\Delta}^\top\rangle) \\
&- \frac{3}{m}\sum_{i=1}^m \langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle\langle A_i, \mathbf{x}\bar{\mathbf{x}}^\top - \mathbf{z}(\bar{\mathbf{z}})^\top\rangle \\
&+ \frac{1}{m}\sum_{i=1}^m \langle A_i, \Delta\bar{\Delta}^\top\rangle\langle A_i, \Delta\bar{\Delta}^\top\rangle + 4\langle \nabla g_\eta(\mathbf{x},\bar{\mathbf{x}}), \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}\rangle \\
&- \frac{2}{m}\sum_{i=1}^m \eta_i \langle A_i, \mathbf{x}\bar{\mathbf{x}} - \mathbf{z}\bar{\mathbf{z}}\rangle - \frac{2}{m}\sum_{i=1}^m \eta_i \langle A_i, \Delta\bar{\mathbf{x}}^T + \mathbf{x}\bar{\Delta}^T\rangle
\end{aligned}
\tag{B.27}
$$

Using Lemma B.4.1 and Lemma B.4.4, we can conclude that with probability greater

than $1 - c_1 e^{-c_2 m \min\{D(t), E(\epsilon)\}}$,

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_0(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix} \\
&\leq -(1-t)\|\Delta\mathbf{x}^T - \mathbf{x}\Delta^T\|_F^2 + 4\delta\|\Delta\|_2 + \beta\|\Delta\Delta^*\|_F^2 \\
&- 3\alpha\|\mathbf{x}\mathbf{x}^* - \mathbf{z}(\mathbf{z})^*\|_F^2 \\
&\leq -(1-t)\|\Delta\mathbf{x}^T - \mathbf{x}\Delta^T\|_F^2 + 4\delta\|\Delta\|_2 \\
&+ 2\beta\|\mathbf{x}\mathbf{x}^* - \mathbf{z}(\mathbf{z})^*\|_F^2 - 3\alpha\|\mathbf{x}\mathbf{x}^* - \mathbf{z}(\mathbf{z})^*\|_F^2 \\
&\leq (2\beta - 3\alpha)\,\|\mathbf{x}\mathbf{x}^* - \mathbf{z}(\mathbf{z})^*\|_F^2 + 4\delta\|\Delta\|_2 \\
&- (1-t)\|\Delta\mathbf{x}^T - \mathbf{x}\Delta^T\|_F^2
\end{aligned} \tag{B.28}$$

where $\exists c_1, c_2 > 0$ which can be computed from Lemma **??** and Lemma B.4.4. Further using bounds from Lemma B.6.1, we can conclude with probability $1 - c_1 e^{-c_2 m \min\{D(t), E(\epsilon)\}} - 2e^{-c_3 m \epsilon_\eta}$,

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_\eta(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix} \\
&\leq (2\beta - 3\alpha)\,\|\mathbf{x}\mathbf{x}^* - \mathbf{z}(\mathbf{z})^*\|_F^2 + 4\delta\|\Delta\|_2 + t \\
&+ 4\epsilon_\eta\sigma_\eta\|\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^*\|_F
\end{aligned}$$

For any $\xi \in [0, 1]$, there can be multiple possibilities of the constants β, ζ and γ satisfying Theorem 3.6.7.

Given ξ, we can bound $m = O(n)$ such that the mapping $\mathcal{M}_\mathcal{A}$ is $(1-c, 1+c)$-stable, for some small $c > 0$ and if the current vector $\mathbf{x}$ is not close to $\mathbf{x}^*$ such that $d(\mathbf{x}, \mathbf{z}) = C_0\delta$, for sufficiently large $C_0 > 0$, then we have

$$\begin{aligned}
&\begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix}^* \nabla^2 g_\eta(\mathbf{x}, \bar{\mathbf{x}}) \begin{bmatrix}\Delta\\ \bar{\Delta}\end{bmatrix} \\
&\leq (-1 + 5c)C_0^2\delta^2 + 4C_0\delta^2 + 4\epsilon_\eta\sigma_\eta C_0\delta
\end{aligned} \tag{B.29}$$

If the above inequality is not satisfied then,

$$\begin{aligned}
(-1 + 5c)C_0\delta + 4\delta + 4\epsilon_\eta\sigma_\eta \;&\geq 0 \\
C_0 \;&\leq \frac{4(\delta + \epsilon_\eta\sigma_\eta)}{\delta(1 - 5c)}
\end{aligned} \tag{B.30}$$

Hence we can bound the distance of the optima of (P1) – $\mathbf{z}$ and the solution of the problem (F_η) – $\hat{\mathbf{x}}$ as,

$$\|\mathbf{x}\mathbf{x}^* - \mathbf{z}\mathbf{z}^*\|_F \leq \frac{4(\delta + \epsilon_\eta\sigma_\eta)}{1 - 5c} \tag{B.31}$$

□

APPENDIX C

MULTI-AGENT BANDIT SEARCH

C.1 Proof for Proposition 4.4.1

Let the undesirable events be $\mathcal{E}_{\mathcal{K}}(t) = \{\mathcal{K}(t) \setminus \mathcal{S}_{\theta-\epsilon} \neq \emptyset\}$ and $\mathcal{E}_{\mathcal{R}}(t) = \left\{\mathcal{R}(t) \setminus \mathcal{S}^{\complement}_{\theta+\epsilon} \neq \emptyset\right\}$. To prove Proposition 4.4.1, we want to show $\mathbb{P}\left[\left(\bigcup_{t\geq 1}\mathcal{E}_{\mathcal{K}}(t)\right) \bigcup \left(\bigcup_{t\geq 1}\mathcal{E}_{\mathcal{R}}(t)\right)\right] \leq \delta$. By Boole's inequality, it is suffices to show that $\mathbb{P}\left[\bigcup_{t\geq 1}\mathcal{E}_{\mathcal{K}}(t)\right] \leq \delta/2$ and $\mathbb{P}\left[\bigcup_{t\geq 1}\mathcal{E}_{\mathcal{R}}(t)\right] \leq \delta/2$.

Recall that, by the choice of $U_{i,\boldsymbol{\pi}}(t,\delta)$ in (4.6c),

$$\mathbb{P}\left[\bigcap_{t\geq 1}\{|\mu_i - \hat{\mu}_i| \leq U_{i,\boldsymbol{\pi}}(t,\delta)\}\right] \geq 1 - \frac{\delta}{2|\mathcal{G}|}, \tag{C.1}$$

for any grid cell $i \in \mathcal{G}$ and search policy $\boldsymbol{\pi}$ (see Lemma 1 in (Jun *et al.*, 2016) with $\omega = \sqrt{\delta/(12|\mathcal{G}|)}$).

For any $t \geq 1$, $\mathcal{K}(t) \setminus \mathcal{S}_{\theta-\epsilon} \neq \emptyset$ if and only if $\hat{\mu}_{i,\boldsymbol{\pi}}(t) - U_{i,\boldsymbol{\pi}}(t,\delta) \geq \theta - \epsilon > \mu_i$ for some $i \in \mathcal{G}$. Consequently, $\mathcal{K}(t) \setminus \mathcal{S}_{\theta-\epsilon} \neq \emptyset \Rightarrow |\hat{\mu}_{i,\boldsymbol{\pi}}(t) - \mu_i| \geq U_{i,\boldsymbol{\pi}}(t,\delta)$ for some $i \in \mathcal{G}$. By (C.1) and Boole's inequality,

$$\begin{aligned}\mathbb{P}\left[\bigcup_{t\geq 1}\mathcal{E}_{\mathcal{K}}(t)\right] &\leq \mathbb{P}\left[\bigcup_{i\in\mathcal{G}}\bigcup_{t\geq 1}\{|\hat{\mu}_{i,\boldsymbol{\pi}}(t) - \mu_i| \geq U_{i,\boldsymbol{\pi}}(t,\delta)\}\right] \\ &\leq \sum_{i\in\mathcal{G}}\frac{\delta}{2|\mathcal{G}|} \leq \frac{\delta}{2}.\end{aligned}$$

The proof for $\mathbb{P}\left[\bigcup_{t\geq 1}\mathcal{E}_{\mathcal{R}}(t)\right] \leq \delta/2$ follows similarly. ■

C.2 Proof Sketch for Theorem 4.5.2

Bounding $T_{\boldsymbol{\pi}}$ (4.10): Recall that the number of sufficient number of samples required for successful classification, with high confidence, of a grid cell $i \in \mathcal{G}$ can be tightly upper-bounded by $\mathcal{O}(\phi_i)$ (see (6) in (Jun *et al.*, 2016) with $\omega = \sqrt{\delta/(2|\mathcal{G}|)}$). If we allowed multiple search agents to visit a grid cell simultaneously, then $\frac{1}{d}\sum_{i\in\mathcal{G}}\mathcal{O}(\phi_i)$ upper bounds $T_{\boldsymbol{\pi}}$, due to the independence assumption between the cells. However, the agents are required to stay in distinct cells at all times. Consequently, some of the search agents are rendered ineffective when less than d cells are left to be labelled. (4.10) upper bounds $T_{\boldsymbol{\pi}}$ by accounting for the worst-case inefficiency — the last d cells is a set of "easy-to-classify" $d-1$ cells and a "hardest-to-classify" cell.

Bounding $L(\boldsymbol{\pi})$ (4.11): We split the time taken to classify all interesting cells by Algorithm 3 into three parts:

i. *Classifying interesting cells* : The number of samples sufficient for classification of interesting cell i is $\mathcal{O}(\phi_i)$,

ii. *Sampling uninteresting cells when biased coin toss yields heads* : Here, Algorithm 3 samples grid cells while maximizing J, see (4.6). The number of sufficient samples of the uninteresting cell j after which, it will be sampled by *only* after classifying *all* interesting cells upper-bounded by $\mathcal{O}(\gamma_j)$. Additionally, we add

a margin of $\frac{4(1-\alpha)|\mathcal{G}|^2}{\alpha\delta}$ to account for the worst-case low-probability event of revisiting cell j due to the switching to label-then-move.

iii. *Sampling uninteresting cells when biased coin toss yields tails* : Here, Algorithm 3 samples grid cells based only on the distance metric ℓ. In the worst case, we may sample an uninteresting cell long enough to classify it, which is $O(\phi_i)$.

Combining these parts, we have

$$L(\boldsymbol{\pi}) \leq \underbrace{\sum \mathcal{O}(\phi_i)}_{\text{interesting cells (i)}} + \underbrace{\sum \min\left\{\mathcal{O}(\phi_i), \mathcal{O}(\gamma_i) + \frac{4(1-\alpha)|\mathcal{G}|^2}{\alpha\delta}\right\}}_{\text{uninteresting cells (ii) and (iii)}}.$$

for a team with $d = 1$. Similar to $T_{\boldsymbol{\pi}}$, we obtain (4.11) by accounting for oversampling due to inefficiency arising from the presence of $d > 1$ agents.

Bounding $E(\boldsymbol{\pi})$ (4.12): The economic cost (4.5) consists of the movement cost and the sampling cost. At every time step t, Algorithm 3 performs an α biased coin-toss. For coin tosses corresponding to heads, Algorithm 3 moves to the grid cells which maximizes the activation function (4.6). In this case, Algorithm 3 incurs a movement cost of at most $M \triangleq \max_{\mathbf{a},\mathbf{a}'} \ell(\mathbf{a}, \mathbf{a}')$. For coin tosses corresponding to tails, Algorithm 3 searches for the nearest unlabelled cell in the neighbourhood. The total cost incurred by Algorithm 3 during the *entire* run is no larger than the cost incurred to visit all of the cells in some pre-defined sequence, which we know is $O(|\mathcal{G}| - 1)$.

$$\sum_{1\leq\tau\leq t} \mathbb{E}\left[\ell(\mathbf{a}_\tau, \mathbf{a}_{\tau-1})\right] \leq (1-\alpha)O(|\mathcal{G}| - 1) + \alpha M t \tag{C.2}$$

The sampling cost accrued at iteration t of Algorithm 3 is $\beta t d$. We complete the proof by adding these bounds, and applying the bound in (4.10) on t. ■