Mathematical Programming
https://doi.org/10.1007/s10107-022-01896-3

FULL LENGTH PAPER

Series A

Solving orthogonal group synchronization via convex and low-rank optimization: tightness and landscape analysis

Shuyang Ling[1]

Received: 1 June 2020 / Accepted: 17 September 2022

Abstract

Group synchronization aims to recover the group elements from their noisy pairwise measurements. It has found many applications in community detection, clock synchronization, and joint alignment problem. This paper focuses on the orthogonal group synchronization which is often used in cryo-EM and computer vision. However, it is generally NP-hard to retrieve the group elements by finding the least squares estimator. In this work, we first study the semidefinite programming (SDP) relaxation of the orthogonal group synchronization and its tightness, i.e., the SDP estimator is exactly equal to the least squares estimator. Moreover, we investigate the performance of the Burer-Monteiro factorization in solving the SDP relaxation by analyzing its corresponding optimization landscape. We provide deterministic sufficient conditions which guarantee: (i) the tightness of SDP relaxation; (ii) optimization landscape arising from the Burer-Monteiro approach is benign, i.e., the global optimum is exactly the least squares estimator and no other spurious local optima exist. Our result provides a solid theoretical justification of why the Burer-Monteiro approach is remarkably efficient and effective in solving the large-scale SDPs arising from orthogonal group synchronization. We perform numerical experiments to complement our theoretical analysis, which gives insights into future research directions.

Keywords Orthogonal group synchronization · Convex optimization · Low-rank optimization · Burer-Monteiro factorization · Optimization landscape

Mathematics Subject Classification 90C22 · 90C26 · 90C46 · 62F10

This work is (partially) financially supported by the National Key R&D Program of China, Project Number 2021YFA1002800, National Natural Science Foundation of China (NSFC) No.12001372, Shanghai Municipal Education Commission (SMEC) via Grant 0920000112, and NYU Shanghai Boost Fund.

✉ Shuyang Ling
sl3635@nyu.edu

[1] New York University Shanghai, 1555 Century Avenue, Shanghai 200122, China

Published online: 01 October 2022

1 Introduction

Group synchronization requires to recover the group elements $\{g_i\}_{i=1}^n$ from their partial pairwise measurements:

$$g_{ij} = g_i^{-1} g_j + w_{ij}, \quad (i, j) \in \mathcal{E}$$

where g_i belongs to a given group $\mathcal{G}$, w_{ij} is the noise, and $\mathcal{E}$ is the edge set of an underlying network. Depending on the specific group choices, one has found many interesting problems including $\mathbb{Z}_2$-synchronization [2], angular synchronization [7, 48], and permutation group [43], special orthogonal group [4, 50], orthogonal group [41], cyclic group [22], and real number addition group [26].

Optimization plays a crucial role in solving group synchronization problems. One commonly used approach is the least squares method. However, the least squares objective arising from many of these aforementioned examples are usually highly nonconvex or even inherently discrete. This has posed a major challenge to retrieve the group elements from their highly noisy pairwise measurements because finding the least squares estimator is NP-hard in general. In the recent few years, many efforts are devoted to finding spectral relaxation and convex relaxation (in particular semidefinite relaxation) as well as nonconvex approaches to solve these otherwise NP-hard problems.

In this work, we focus on the general orthogonal synchronization problem:

$$\boldsymbol{A}_{ij} = \boldsymbol{G}_i \boldsymbol{G}_j^\top + \text{noise}$$

where $\boldsymbol{G}_i$ is a $d \times d$ orthogonal matrix belonging to

$$\mathrm{O}(d) := \left\{ \boldsymbol{O} \in \mathbb{R}^{d\times d} : \boldsymbol{O}^\top \boldsymbol{O} = \boldsymbol{O}\boldsymbol{O}^\top = \boldsymbol{I}_d \right\}. \tag{1}$$

Orthogonal group synchronization is frequently found in cryo-EM [50], computer vision [4] and feature matching problem [26], and is a natural generalization of $\mathbb{Z}_2$- and angular synchronization [2, 48]. We aim to establish a theoretical framework to understand when convex and nonconvex approaches can retrieve the orthogonal group elements from the noisy measurements. In particular, we will answer the following two core questions.

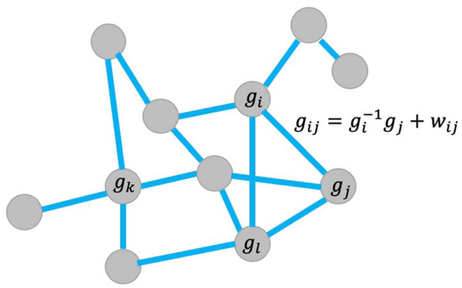

Fig. 1 Illustration of group synchronization on networks

Table 1 Examples of group synchronization and applications

Group	Application
$\mathbb{Z}_2 = \{1, -1\}$	Community detection and node label recovery [2, 14]
Cyclic group $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$	Joint alignment from pairwise differences [22]
$U(1) = \{e^{\mathrm{i}\theta} : \theta \in [0, 2\pi)\}$	Angular synchronization [7, 48, 63], Phase retrieval [29]
Permutation group	Feature matching [28, 43]
SO(3): special orthogonal group	Cryo-EM [49, 50] and computer vision [4]
Addition group on a finite interval	Clock synchronization on networks [26]
SO(d) or O(d)	Procrustes problem [4, 21, 33], Rotation averaging [23]
Special Euclidean group	Simultaneous localization and mapping [47]

- For the convex relaxation of orthogonal group synchronization,

  When is the convex relaxation tight?

  Convex relaxation has proven to be extremely powerful in approximating or giving the exact solution to many otherwise NP-hard problems under certain conditions. However, it is not always the case that the relaxation yields a solution that matches the least squares estimator. Thus we aim to answer when one can find the least squares estimator of O(d) synchronization with a simple convex program.
- For the nonconvex Burer-Monteiro approach [16, 17], we are interested in answering this question:

  When does the Burer-Monteiro approach yield a benign optimization landscape?

  Empirical evidence has indicated that the Burer-Monteiro approach works extremely well to solve large-scale SDPs in many applications despite its inherent nonconvexity. One way to provide a theoretical justification is to show that the optimization landscape is *benign*, i.e., there is only one global minimizer and no other *spurious* local minima exist.

1.1 Related works and our contribution

Group synchronization problem is a rich source of many mathematical problems. Now we will give a review of the related works which inspire this work. Table 1 provides a non-exhaustive summary of important examples in group synchronization with applications.

Various approaches have been proposed to solve the group synchronization problem [2, 32, 43, 45, 48, 63]. One of the most commonly used approaches in group synchronization is to find the least squares estimator. As pointed out earlier, finding the least squares estimator of general O(d) synchronization is NP-hard. In fact, if the group is $\mathbb{Z}_2$, the least squares objective function is closely related to the graph MaxCut problem, which is a well-known NP-hard problem. Therefore, alternative methods are often needed to deal with this situation. One line of research focuses on the spectral and semidefinite programming relaxation (SDP) including $\mathbb{Z}_2$-synchronization [2], angular synchronization [7, 48, 63], orthogonal group [4, 21, 35, 50, 57], and permutation group [28, 43]. Regarding the SDP relaxation, many efforts are devoted to designing approximation algorithms, such as the Goemans-Williamson relaxation [27] for graph MaxCut. Inspired by the development of compressive sensing and low-rank recovery [19, 46], we are more interested in the tightness of convex relaxation: the data are not as adversarial as expected in some seemingly NP-hard problems. Convex relaxation in these problems admits the exact solution to the original seemingly NP-hard problem under certain conditions. Following this idea, [2] studies the SDP relaxation of $\mathbb{Z}_2$-synchronization with corrupted data; the tightness of the SDP relaxation for angular synchronization is first investigated in [7], followed by works [13, 37], and a near-optimal performance bound is obtained in [63]. The work [57] studies the convex relaxation of $\mathrm{SO}(d)$ group synchronization with uniform random corruption (multiplicative noise) and successfully characterizes the phase transition for the exact recovery of the group elements. A very recent work [62] gives a sufficient condition to ensure the tightness of the SDP for the orthogonal trace-sum maximization problem, which includes orthogonal group synchronization as a special case. The condition in [62] is similar to the one presented in this work but differs from ours in the dimension factor of the group element.

Despite the effectiveness of convex approaches and recent developments of excellent general SDP solvers such as SDPNAL+ [51, 60] and SDPT3 [54], it is still highly expensive to solve large-scale SDPs [42]. It is more advantageous to keep the low-rank structure and obtain a much more efficient algorithm [3, 12, 61]: orthogonal group synchronization can be first formulated as a low-rank optimization program with orthogonality constraints and then tackled with many general solvers [12, 24, 58]. As a result, there is a growing interest in developing nonconvex approaches, particularly the first-order gradient-based algorithm. These methods enjoy the advantage of higher efficiency than the convex approach. However, there are also concerns about the possible existence of multiple local optima which prevent the iteration from converging to the global one. The recent few years have witnessed a surge of research in exploring fast and provably convergent nonconvex optimization approaches. Two main strategies are: (i) design a smart initialization scheme and then provide the global convergence; (ii) analyze the nonconvex optimization landscape. Examples include phase retrieval [18, 53], dictionary learning [52], joint alignment [22], matrix completion [25, 31] and spiked tensor model [9]. These ideas are also applied to several group synchronization problems. For example, the works [22] on joint alignment, [13, 37, 63] on angular synchronization, and [34, 35, 38] on orthogonal group synchronization follow the two-step procedures: first, use the spectral method for a good initialization and then show that the projected generalized power methods have the property of global linear convergence.

Our focus here is on the optimization landscape of the Burer-Monteiro approach [16, 17] in solving the large SDPs arising from $\mathrm{O}(d)$ synchronization. The original remarkable work [17] by Burer and Monteiro shows that as long as $p(p+1) > 2n$ where p is the dimension of low-rank matrix and n is the number of constraints, the global optima to Burer-Monteiro factorization match those of the corresponding SDP by using the idea from [44]. Later on, [14, 15] show that the optimization landscape is benign, meaning that no spurious local optima exist in the nonconvex objective function if p is approximately greater than $\sqrt{2n}$. This bound is proven to be almost tight in [56]. On the other hand, it is widely believed that even if $p = O(1)$, the Burer-Monteiro factorization works provably, which is supported by many numerical experiments. We have benefitted greatly from the works regarding the Burer-Monteiro approach on group synchronization in [8, 41, 59]. In [8], the authors prove that the optimization landscape is benign for $\mathbb{Z}_2$-synchronization as well as community detection under the stochastic block model if $p = 2$. The optimization landscape of angular synchronization is studied in [59]. The work [41] provides a lower bound for the objective function value evaluated at local optima. The bound depends on the rank p and is smartly derived by using the Riemannian Hessian. However, the landscape and the tightness of the Burer-Monteiro approach for $\mathrm{O}(d)$ synchronization have not been fully addressed yet, which becomes one main motivation for this work. It is worth noting that the Burer-Monteiro approach is closely related to the synchronization of oscillators on manifold [36, 39, 40]. The analysis of the optimization landscape of the Burer-Monteiro approach in $\mathbb{Z}_2$ with $p = 2$ is equivalent to exploring the stable equilibria of the energy landscape associated with the homogeneous Kuramoto oscillators [36]. This connection is also reflected in the synchronization of coupled oscillators on more general manifolds such as n-sphere and Stiefel manifold [39, 40] on arbitrary complex networks.

An important problem regarding the tightness of convex relaxation and landscape analysis is how these two properties depend on the general notion of SNR (signal-to-noise ratio). In most cases, if the noise is rather small compared to the planted signal, optimization methods should easily recover the hidden signal since the tightness of SDP and the benign landscape are guaranteed. However, as the noise strengthens, the landscape becomes bumpy and optimizing the cost function becomes challenging. This leads to the research topic on detecting the critical threshold for this phase transition. Examples can be found in many applications including eigenvectors estimation [5] and the community detection under the stochastic block model [1]. For $\mathbb{Z}_2$- and angular synchronization, convex methods are tight all the way to the information-theoretical limit [6, 63] but the analysis of optimization landscape remains suboptimal [8, 59]. Our work on $\mathrm{O}(d)$ synchronization will follow a similar idea and attempt to explore the critical threshold to ensure the tightness of convex relaxation and the Burer-Monteiro approach.

Our contribution is multifold. First, we prove the tightness of convex relaxation in solving the $\mathrm{O}(d)$ synchronization problem by extending the work [7, 8] on $\mathbb{Z}_2$- and angular synchronization. We propose a deterministic sufficient condition that guarantees the tightness of the SDP relaxation and easily applies to other noise models. Our result slightly improves the very recent result on the tightness of the $\mathrm{O}(d)$ synchronization in [62]. Moreover, we analyze the optimization landscape arising from the Burer-Monteiro approach applied to $\mathrm{O}(d)$ synchronization, which has not been investigated in the aforementioned works. For this low-rank optimization approach, we also provide a general deterministic condition to ensure a benign optimization landscape. The sufficient condition is quite general and applicable to several aforementioned examples including community detection under the stochastic block model [1, 8], $\mathbb{Z}_2$- and angular synchronization [8, 13, 37, 59], and $\mathrm{SO}(d)$ synchronization under uniform corruption [57], and achieves the state-of-the-art results on the analysis of optimization landscape. Our result on the landscape analysis serves as another example to demonstrate the great success of the Burer-Monteiro approaches in solving large-scale SDPs.

1.2 Organization of this paper

The paper proceeds as follows: Sect. 2 introduces the background of orthogonal group synchronization and optimization methods. We show the main results in Sect. 3. Section 4 focuses on numerical experiments and we give the proofs in Sect. 5.

1.3 Notation

For any given matrix $\boldsymbol{X}$, $\boldsymbol{X}^\top$ is the transpose of $\boldsymbol{X}$; $\boldsymbol{X} \succeq 0$ means $\boldsymbol{X}$ is positive semidefinite. Denote $\|\boldsymbol{X}\|_{\mathrm{op}}$ the operator norm of $\boldsymbol{X}$, $\|\boldsymbol{X}\|_F$ the Frobenius norm, and $\|\boldsymbol{X}\|_*$ the nuclear norm, i.e., the sum of singular values. For two matrices $\boldsymbol{X}$ and $\boldsymbol{Z}$ of the same size, $\boldsymbol{X} \circ \boldsymbol{Z}$ denotes the Hadamard product of $\boldsymbol{X}$ and $\boldsymbol{Z}$, i.e., $(\boldsymbol{X} \circ \boldsymbol{Z})_{ij} = X_{ij} Z_{ij}$; $\langle \boldsymbol{X}, \boldsymbol{Z} \rangle := \mathrm{Tr}(\boldsymbol{X}\boldsymbol{Z}^\top)$ is the inner product. "$\otimes$" stands for the Kronecker product; $\mathrm{diag}(\boldsymbol{v})$ gives a diagonal matrix whose diagonal entries equal $\boldsymbol{v}$; $\mathrm{blkdiag}(\boldsymbol{\Pi}_{11}, \cdots, \boldsymbol{\Pi}_{nn})$ denotes a block-diagonal matrix whose diagonal blocks are $\boldsymbol{\Pi}_{ii}$, $1 \leq i \leq n$. Let $\boldsymbol{I}_n$ be the $n \times n$ identity matrix, and $\boldsymbol{J}_n$ be an $n \times n$ matrix whose entries are 1. We denote $\boldsymbol{X} \succeq \boldsymbol{Z}$ if $\boldsymbol{X} - \boldsymbol{Z} \succeq 0$, i.e., $\boldsymbol{X} - \boldsymbol{Z}$ is positive semidefinite. We write $f(n) \lesssim g(n)$ for two positive functions $f(n)$ and $g(n)$ if there exists an absolute positive constant C such that $f(n) \leq C g(n)$ for all n.

2 Preliminaries

2.1 The model of group synchronization

We introduce the model for orthogonal group synchronization. We want to estimate n matrices $\boldsymbol{G}_1, \cdots, \boldsymbol{G}_n \in \mathrm{O}(d)$ from their pairwise measurements $\boldsymbol{A}_{ij}$:

$$\boldsymbol{A}_{ij} = \boldsymbol{G}_i \boldsymbol{G}_j^{-1} + \boldsymbol{\Delta}_{ij}, \tag{2}$$

where $\boldsymbol{A}_{ij} \in \mathbb{R}^{d\times d}$ is the observed data and $\boldsymbol{\Delta}_{ij} \in \mathbb{R}^{d\times d}$ is the additive noise. Note that $\boldsymbol{G}_i^{-1} = \boldsymbol{G}_i^\top$ holds for any orthogonal matrix $\boldsymbol{G}_i$. Thus in the matrix form, we can reformulate the observed data $\boldsymbol{A}$ as

$$\boldsymbol{A} = \boldsymbol{G}\boldsymbol{G}^\top + \boldsymbol{\Delta} \in \mathbb{R}^{nd\times nd}$$

where $\boldsymbol{G}^\top = [\boldsymbol{G}_1^\top, \cdots, \boldsymbol{G}_n^\top] \in \mathbb{R}^{d\times nd}$ and the (i, j)-block of $\boldsymbol{A}$ is $\boldsymbol{A}_{ij} = \boldsymbol{G}_i \boldsymbol{G}_j^\top + \boldsymbol{\Delta}_{ij}$. In particular, we set $\boldsymbol{A}_{ii} = \boldsymbol{I}_d$ and $\boldsymbol{\Delta}_{ii} = 0$. The model (2) includes many interesting applications as special cases including $\mathbb{Z}_2$- and angular synchronization [2, 13, 37, 48, 57, 63], community detection [1, 8], and $\mathrm{SO}(d)$ synchronization [57]. In this work, we will establish our theory for the general model (2) and apply it to one specific model, i.e., the $\mathrm{O}(d)$ synchronization under additive Gaussian noise. Our main results also apply to these aforementioned examples but we will not provide all the technical details here.

One benchmark noise model is the group synchronization from measurements corrupted with Gaussian noise, i.e.,

$$\boldsymbol{A}_{ij} = \boldsymbol{G}_i \boldsymbol{G}_j^\top + \sigma \boldsymbol{W}_{ij}$$

where each entry in $\boldsymbol{W}_{ij}$ is an i.i.d. standard Gaussian random variable and $\boldsymbol{W}_{ij}^\top = \boldsymbol{W}_{ji}$. The corresponding matrix form is

$$\boldsymbol{A} = \boldsymbol{G}\boldsymbol{G}^\top + \sigma \boldsymbol{W} \in \mathbb{R}^{nd\times nd}$$

which is actually a matrix spike model.

One common approach to recover $\boldsymbol{G}$ is to minimize the nonlinear least squares objective function over the orthogonal group $\mathrm{O}(d)$:

$$\min_{\boldsymbol{R}_i\in \mathrm{O}(d)} \sum_{i=1}^{n}\sum_{j=1}^{n} \left\| \boldsymbol{R}_i\boldsymbol{R}_j^\top - \boldsymbol{A}_{ij}\right\|_F^2 \tag{3}$$

which is also known as the rotation averaging problem [23]. In fact, the global minimizer equals the maximum likelihood estimator of (2) under Gaussian noise, i.e., assuming each $\boldsymbol{\Delta}_{ij}$ is an independent Gaussian random matrix.

Throughout our discussion, we will deal with a more convenient equivalent form. More precisely, we perform a change of variable:

$$\begin{aligned}\left\| \boldsymbol{R}_i\boldsymbol{R}_j^\top - \boldsymbol{A}_{ij}\right\|_F^2 &= \left\| \boldsymbol{R}_i\boldsymbol{R}_j^\top - (\boldsymbol{G}_i\boldsymbol{G}_j^\top + \boldsymbol{\Delta}_{ij})\right\|_F^2 \\ &= \left\| \boldsymbol{G}_i^\top\boldsymbol{R}_i\boldsymbol{R}_j^\top\boldsymbol{G}_j - \left(\boldsymbol{I}_d + \boldsymbol{G}_i^\top\boldsymbol{\Delta}_{ij}\boldsymbol{G}_j\right)\right\|_F^2\end{aligned}$$

where $\boldsymbol{R}_i$ and $\boldsymbol{G}_i \in \mathrm{O}(d)$. By letting

$$\boldsymbol{R}_i \leftarrow \boldsymbol{G}_i^\top \boldsymbol{R}_i, \quad \boldsymbol{\Delta}_{ij} \leftarrow \boldsymbol{G}_i^\top\boldsymbol{\Delta}_{ij}\boldsymbol{G}_j \tag{4}$$

then the updated objective function becomes

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\left\| \boldsymbol{R}_i\boldsymbol{R}_j^\top - \left(\boldsymbol{I}_d + \boldsymbol{\Delta}_{ij}\right)\right\|_F^2.$$

Its global minimizer equals the global maximizer to

$$\max_{\boldsymbol{R}_i\in \mathrm{O}(d)} \sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle \boldsymbol{R}_i\boldsymbol{R}_j^\top, \boldsymbol{I}_d + \boldsymbol{\Delta}_{ij}\right\rangle \tag{P}$$

The program (P) is a well-known NP-hard problem. We will focus on solving (P) by convex relaxation and low-rank optimization approach, and study their theoretical guarantees.

2.2 Convex relaxation

The convex relaxation relies on the idea of lifting: let $\boldsymbol{X} = \boldsymbol{R}\boldsymbol{R}^\top \in \mathbb{R}^{nd\times nd}$ with $\boldsymbol{X}_{ij} = \boldsymbol{R}_i\boldsymbol{R}_j^\top$. We notice that $\boldsymbol{X} \succeq 0$ and $\boldsymbol{X}_{ii} = \boldsymbol{I}_d$ hold for any $\{\boldsymbol{R}_i\}_{i=1}^n \in \mathrm{O}(d)$. The convex relaxation of $\mathrm{O}(d)$ synchronization is

$$\max_{\boldsymbol{X}\in\mathbb{R}^{nd\times nd}} \langle \boldsymbol{A}, \boldsymbol{X}\rangle \quad \text{such that} \quad \boldsymbol{X}_{ii} = \boldsymbol{I}_d, \quad \boldsymbol{X} \succeq 0 \tag{SDP}$$

where (i, j)-block of $\boldsymbol{A}$ is $\boldsymbol{A}_{ij} = \boldsymbol{I}_d + \boldsymbol{\Delta}_{ij} \in \mathbb{R}^{d\times d}$. In particular, if $d = 1$, this semidefinite programming (SDP) relaxation reduces to the famous Goemans-Williamson relaxation for the graph MaxCut problem [27]. Since we relax the constraint, it is not necessarily the case that the global maximizer $\widehat{\boldsymbol{X}}$ to (SDP) is exactly rank-d, i.e., $\widehat{\boldsymbol{X}} = \widehat{\boldsymbol{G}}\widehat{\boldsymbol{G}}^\top$ for some $\widehat{\boldsymbol{G}} \in \mathbb{R}^{nd\times d}$ with $\widehat{\boldsymbol{G}}_i \in \mathrm{O}(d)$. Our goal is to study *the tightness of this SDP relaxation*: when the solution to (SDP) is exactly rank-d, i.e., the convex relaxation gives the global optimal solution to (P) which is also the least squares estimator.

2.3 Low-rank optimization: Burer-Monteiro approach

Note that solving the convex relaxation (SDP) is extremely expensive especially for large d and n. Thus an efficient, robust, and provably convergent optimization algorithm is always in great need. Since the solution is usually low-rank in many empirical experiments, it is appealing to take advantage of this property: keep the iterates low-rank and perform first-order gradient-based approach to solve this otherwise computationally expensive SDP. In particular, we will resort to the Burer-Monteiro approach [16, 17] to deal with the orthogonal group synchronization problem. The core idea of the Burer-Monteiro approach is keeping $\boldsymbol{X}$ in a factorized form and taking advantage of its low-rank property. Recall the constraints in (SDP) read $\boldsymbol{X} \succeq 0$ and $\boldsymbol{X}_{ii} = \boldsymbol{I}_d$. In the Burer-Monteiro approach, we let $\boldsymbol{X} = \boldsymbol{S}\boldsymbol{S}^\top$ where $\boldsymbol{S} \in \mathbb{R}^{nd\times p}$ with $p > d$. We hope to recover the group elements by maximizing $f(\boldsymbol{S})$:

$$f(\boldsymbol{S}) := \langle \boldsymbol{A}, \boldsymbol{S}\boldsymbol{S}^\top\rangle = \sum_{i=1}^n\sum_{j=1}^n \left\langle \boldsymbol{A}_{ij}, \boldsymbol{S}_i\boldsymbol{S}_j^\top \right\rangle, \tag{BM}$$

where

$$\boldsymbol{S}^\top := \left[\boldsymbol{S}_1^\top, \cdots, \boldsymbol{S}_n^\top\right] \in \mathbb{R}^{p\times nd}, \quad \boldsymbol{S}_i\boldsymbol{S}_i^\top = \boldsymbol{I}_d.$$

In other words, we substitute $\boldsymbol{R}_i$ in (P) by a partial orthogonal matrix $\boldsymbol{S}_i \in \mathbb{R}^{d\times p}$ with $\boldsymbol{S}_i\boldsymbol{S}_i^\top = \boldsymbol{I}_d$ and $p > d$. Therefore, $\boldsymbol{S}$ belongs to the product space of Stiefel manifold, i.e., $\mathrm{St}(d,p)^{\otimes n} := \underbrace{\mathrm{St}(d,p)\times\cdots\times\mathrm{St}(d,p)}_{n\text{ times}}$

$$\mathrm{St}(d,p) := \left\{\boldsymbol{S}_i \in \mathbb{R}^{d\times p} : \boldsymbol{S}_i\boldsymbol{S}_i^\top = \boldsymbol{I}_d\right\}. \tag{5}$$

Running projected gradient method on this objective function (BM) definitely saves a large amount of computational resources and memory storage. However, the major issue here is the nonconvexity[1] of the objective function, i.e., there may exist multiple local maximizers in (BM) and random initialization may lead to one of the local maximizers instead of converging to the global one. As a result, our second focus of this paper is to understand when the Burer-Monteiro approach works for $\mathrm{O}(d)$ synchronization. In particular, we are interested in the optimization landscape of $f(\boldsymbol{S})$: *when does there exist only one global maximizer, without any other spurious local maximizers? Moreover, is this global maximizer exactly rank-d, i.e., it matches the solution to (P)?*

3 Main theorem

3.1 Tightness of SDP relaxation

Here is our main theorem which provides a deterministic condition to ensure the tightness of SDP relaxation in orthogonal group synchronization.

Theorem 1 (Deterministic condition for the tightness of SDP relaxation) *The solutions to (P) and (SDP) are exactly the same, i.e., the global maximizer to the SDP is unique and exactly rank-d, if*

$$n \geq \frac{3\delta^2 d\|\boldsymbol{\Delta}\|_{\mathrm{op}}^2}{2n} + \delta\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}} \max_{1\leq i\leq n}\|\boldsymbol{\Delta}_i\|_{\mathrm{op}} + \max_{1\leq i\leq n}\left\|\boldsymbol{\Delta}_i^\top\boldsymbol{G}\right\|_{\mathrm{op}} + \|\boldsymbol{\Delta}\|_{\mathrm{op}}, \quad \delta = 4, \tag{6}$$

where $\boldsymbol{\Delta}_i^\top = [\boldsymbol{\Delta}_{i1},\cdots,\boldsymbol{\Delta}_{in}] \in \mathbb{R}^{d\times nd}$ *is the ith block row of* $\boldsymbol{\Delta}$ *and* $\boldsymbol{\Delta}_i^\top\boldsymbol{G} = \sum_{j=1}^n \boldsymbol{\Delta}_{ij}\boldsymbol{G}_j$.

The condition in Theorem 1 is purely deterministic and thus it can be easily applied to $\mathrm{O}(d)$ synchronization under other noise models. Theorem 1 indicates that solving the SDP relaxation yields the global maximizer to (P) which is NP-hard in general, under the condition that the noise strength $\|\boldsymbol{\Delta}\|_{\mathrm{op}}$ is small. More precisely, under many statistical models (e.g. for additive Gaussian noise or uniform corruption model), the leading term on the right hand of (6) is given by its first two terms. For simplicity, we may write the condition (6) into

$$n \gtrsim \sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}}^2 \Longleftrightarrow \|\boldsymbol{\Delta}\|_{\mathrm{op}} \lesssim \frac{n^{3/4}}{d^{1/4}} \Longleftrightarrow n \gtrsim d^{1/3}\|\boldsymbol{\Delta}\|_{\mathrm{op}}^{4/3}$$

[1] Here we are actually referring to the nonconvexity of $-f(\boldsymbol{S})$.

where $\max_{1\le i\le n}\|\boldsymbol{\Delta}_i\|_{\mathrm{op}} \le \|\boldsymbol{\Delta}\|_{\mathrm{op}}$.

Here we provide one such example under Gaussian random noise.

Theorem 2 (Tightness of recovery under Gaussian noise) *For $\boldsymbol{\Delta} = \sigma \boldsymbol{W}$ where $\boldsymbol{W}$ is an $nd \times nd$ Gaussian random matrix, the solution to the SDP relaxation (SDP) is unique and exactly rank-d with high probability if*

$$\sigma \le \frac{C_0 n^{1/4}}{d^{3/4}}$$

for some small constant C_0.

Theorem 2 naturally generalizes the work [7] on the tightness of SDP relaxation for angular synchronization under additive Gaussian noise. While the proof of Theorem 2 follows from the idea in [7], several technical parts are different from [7] since $\mathrm{O}(d)$ group is no longer commutative, unlike $U(1)$ group (corresponding to angular synchronization). We will highlight the technical differences in the proof section. Our result improves the bound on σ by a factor of $\sqrt{d}$, compared with the recent result by Zhang [62] in which the tightness of the SDP holds with high probability if

$$\sigma \le \frac{n^{1/4}}{16 d^{5/4}}.$$

To make the constant C_0 more explicit in Theorem 2, we may set $\|\boldsymbol{W}\|_{\mathrm{op}} \le 3\sqrt{nd}$ and $\max_{1\le i\le n}\|\boldsymbol{W}_i^\top \boldsymbol{G}\|_{\mathrm{op}} \le 3\sqrt{n}(\sqrt{d}+\sqrt{2\log n})$ which hold with high probability, and substitute them into (6). Then it suffices to require $\sigma \le n^{1/4}/(300 d^{3/4})$ since only the first two terms in (6) dominate.

One natural question is whether the bound shown (6) above is optimal. The answer is negative, at least under certain statistical models. Take the case with Gaussian noise as an example: the data matrix $\boldsymbol{A} = \boldsymbol{G}\boldsymbol{G}^\top + \sigma \boldsymbol{W}$ is in the signal-plus-noise form. For this spiked matrix model, the strength of the planted signal $\boldsymbol{G}\boldsymbol{G}^\top$ is $\|\boldsymbol{G}\boldsymbol{G}^\top\|_{\mathrm{op}} = n$. The operator norm of the noise is

$$\|\boldsymbol{\Delta}\|_{\mathrm{op}} = \sigma\|\boldsymbol{W}\|_{\mathrm{op}} = 2\sigma\sqrt{nd}(1+o(1))$$

where $\|\boldsymbol{W}\|_{\mathrm{op}} = 2\sqrt{nd}(1+o(1))$ is a classical result for symmetric Gaussian random matrix. In random matrix theory, many researches have been done to identify the detection threshold of the planted signal (a low-rank matrix) [11, 20, 30, 45]. We expect that the hidden signal $\boldsymbol{G}$ can be recovered if the signal strength $\|\boldsymbol{G}\boldsymbol{G}^\top\|_{\mathrm{op}}$ is stronger than the noise $\|\boldsymbol{W}\|_{\mathrm{op}}$, i.e., $\sigma \lesssim \sqrt{n/d}$. Though not explicitly stated, it is believed that $\sigma = \sqrt{n/d}$ is the threshold above which is information-theoretically impossible to detect the signals [30].

In particular, if $d = 1$, i.e., $\mathbb{Z}_2$-synchronization, the SDP relaxation is proven to be tight in [6] with high probability if $\sigma < \sqrt{\frac{n}{(2+\epsilon)\log n}}$ under additive Gaussian noise. If $d = 2$ and

$$W_{ij} = \frac{1}{\sqrt{2}} \begin{bmatrix} X_{ij} & Y_{ij} \\ -Y_{ij} & X_{ij} \end{bmatrix}$$

where X_{ij} and Y_{ij} are independent standard normal, then the model is equivalent to the angular synchronization under complex normal noise which is discussed in [7, 63]. It is shown in [63] that the factor $n^{1/4}$ can be improved to $n^{1/2}$ by using the leave-one-out technique.

For general d, our numerical experiments in Sect. 4 confirm the threshold of order $O(\sqrt{n/d})$: for additive Gaussian noise, the SDP relaxation is tight with high probability if $\sigma \leq c_0\sqrt{n/d}$ for some small constant c_0. In fact, after the completion of this work, the author has proven in [34], by using leave-one-out technique in [63], that when $\sigma \lesssim \sqrt{n}/d$, the SDP relaxation is tight with high probability for additive Gaussian noise $\boldsymbol{\Delta} = \sigma \boldsymbol{W}$ which differs from the information-theoretic threshold only by a factor of $\sqrt{d}$. Therefore, our current analysis still has a large room for improvement, especially for general noise matrix, which we leave for future work.

3.2 Optimization landscape of Burer-Monteiro approach

Our second main result characterizes the optimization landscape of (BM).

Theorem 3 (Uniqueness and tightness of local maximizer) *For the objective function $f(\boldsymbol{S})$ defined in (BM), it has a unique local maximizer (modulo a global orthogonal group action) which is also the global maximizer if $p \geq 2d+1$ and*

$$n \geq \frac{3\delta^2 d\|\boldsymbol{\Delta}\|_{\mathrm{op}}^2}{2n} + \delta\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}} \max_{1\leq i\leq n} \|\boldsymbol{\Delta}_i\|_{\mathrm{op}} + \max_{1\leq i\leq n} \left\|\boldsymbol{\Delta}_i^\top \boldsymbol{G}\right\|_{\mathrm{op}} + \|\boldsymbol{\Delta}\|_{\mathrm{op}} \quad (7)$$

where

$$\delta = \frac{\left(2+\sqrt{5}\right)(p+d)\gamma}{p-2d}, \quad \gamma := \frac{\|\operatorname{Tr}_d(\boldsymbol{\Delta})\|_{\mathrm{op}}}{\|\boldsymbol{\Delta}\|_{\mathrm{op}}} \vee 1 \quad (8)$$

and $\operatorname{Tr}_d(\boldsymbol{\Delta}) = [\operatorname{Tr}(\boldsymbol{\Delta}_{ij})]_{ij} \in \mathbb{R}^{n\times n}$ denotes the partial trace of $\boldsymbol{\Delta}$. Moreover, this global maximizer is exactly rank-d.

Theorem 3 conveys two messages: one is that the optimization landscape of (BM) is benign, meaning there exists a unique local maximizer $\boldsymbol{S} \in \mathbb{R}^{nd\times p}$ (modulo a global orthogonal group action) and $\boldsymbol{S}\boldsymbol{S}^\top$ corresponds to the global maximizer to the SDP relaxation; moreover, it is rank-d, indicating the tightness of the global maximizer. The characterization of benign optimization landscape justifies the remarkable performance of the Burer-Monteiro approach. Moreover, the condition (7), similar to (6), is completely deterministic and does not require any assumption on the statistical properties of the noise, which allows its applications in different settings.

Here are a few remarks on Theorem 3 and its relation to Theorem 1. We start with the choice of p: one always wants to keep p as small as possible since it means cheaper computational costs per iteration. Theorem 3 only requires $p \geq 2d+1$ which is a significant improvement from the generic bound of the Burer-Monteiro approach $p(p+1) > 2nd$ for a benign optimization landscape in [15, 17, 44, 56]. The condition $p \geq 2d+1$ comes from the characterization of the second order critical points of $f(\boldsymbol{S})$, see Lemma 6. We believe the bound can be improved to $p \geq 2d$ or even to $p \geq d+1$ from our current bound $p \geq 2d+1$ with more careful analyses. In our numerical experiments, we have seen that $p = 2d$ suffices to ensure global convergence of the generalized power method from any random initialization.

On the other hand, Theorem 3 indicates that the optimal solution must be exactly rank-d. One may wonder why not choosing $p = d$? Indeed, setting $p = d$ suffices to produce a satisfactory solution by using generalized power method [34, 35, 38] with a proper initialization such as spectral initialization. However, it is a quite different story for the analysis of the optimization landscape of $f(\boldsymbol{S})$. Intuitively, we need to have at least $p \geq d+1$ since $\mathrm{O}(d)$ has two connected components (the determinant of an orthogonal matrix is either 1 or -1). Therefore, the whole manifold $\mathrm{St}(d,p)^{\otimes n}$ is disconnected and has 2^n connected components if $p = d$. Running gradient method on the manifold might not enable the iterates on one connected component to jump to the other one provided that the iterates are moving continuously on the manifold. It is not immediately clear how to guarantee a benign landscape for the objective function defined on such a disconnected manifold.

By comparing (7) with (6), we can see the condition (7) in Theorem 3 only differs from that in Theorem 1 in the parameter δ. Here δ in (7) depends on the dimension p and d of $\mathrm{St}(d,p)$, the partial trace $\mathrm{Tr}_d(\boldsymbol{\Delta})$, and the operator norm of the noise matrix, unlike the constant δ in (6). This difference results from the technical proof in showing the proximity condition of all second-order critical points, i.e., all the second order critical points are highly aligned with the ground truth signal, see Proposition 3 and 4 for details.

Now we apply Theorem 3 to $\mathrm{O}(d)$ synchronization with Gaussian noise.

Theorem 4 (Optimization landscape under Gaussian noise) *The optimization landscape of $f(\boldsymbol{S})$ is benign if $p \geq 2d+1$ and*

$$\sigma \leq C_0 \sqrt{\frac{p-2d}{p+d}} \cdot \frac{n^{1/4}}{d^{3/4}}$$

with high probability for some small constant C_0. In other words, $f(\boldsymbol{S})$ in (BM) has only one local maximizer which is also global and corresponds to the maximizer of the SDP relaxation (SDP) and (P). $\boldsymbol{S}$ is a global maximizer (modulo a global orthogonal transform) and $\boldsymbol{S}\boldsymbol{S}^\top$ is the unique global maximizer to (SDP).

Similar to the scenario in Theorem 2, this bound is likely to be suboptimal in n for additive Gaussian noise. The numerical experiments indicate that $\sigma < 2^{-1}n^{1/2}d^{-1/2}$ should be the optimal scaling. However, it remains one major open problem to prove that the landscape is benign for σ up to the order $n^{1/2}$, even in the scenario of the angular synchronization [8, 36, 37, 59, 63]. We also leave the characterization of the optimization landscape to the future research agenda.

4 Numerics

4.1 Synchronization with Gaussian noise

Our first experiment is to test how the tightness of (SDP) depends on the noise strength. Consider the group synchronization problem,

$$\boldsymbol{A}_{ij} = \boldsymbol{I}_d + \sigma \boldsymbol{W}_{ij} \in \mathbb{R}^{d\times d}, \quad \boldsymbol{A} = \boldsymbol{Z}\boldsymbol{Z}^\top + \sigma \boldsymbol{W}$$

where $\boldsymbol{Z}^\top = [\boldsymbol{I}_d, \cdots, \boldsymbol{I}_d] \in \mathbb{R}^{d\times nd}$ and $\boldsymbol{W} \in \mathbb{R}^{nd\times nd}$ is a symmetric Gaussian random matrix. Since solving (SDP) is rather expensive, we will take an alternative way to find the SDP solution. We first use thd projected power method to get a candidate solution and then confirm it is the global maximizer of the SDP relaxation (SDP) by verifying the global optimality condition.

Proposition 1 indicates that $\widehat{\boldsymbol{R}}\widehat{\boldsymbol{R}}^\top \in \mathbb{R}^{nd\times nd}$ with $\widehat{\boldsymbol{R}} \in \mathrm{O}(d)^{\otimes n}$ is the unique global optimal solution to (P) and (SDP) if

$$\left(\widehat{\boldsymbol{\Lambda}} - \boldsymbol{A}\right)\widehat{\boldsymbol{R}} = 0, \quad \lambda_{d+1}\left(\widehat{\boldsymbol{\Lambda}} - \boldsymbol{A}\right) > 0$$

where $\widehat{\boldsymbol{\Lambda}}$ is an $nd \times nd$ block-diagonal matrix with its ith block equal to

$$\widehat{\boldsymbol{\Lambda}}_{ii} = \frac{1}{2}\sum_{j=1}^{n}\left(\widehat{\boldsymbol{R}}_i\widehat{\boldsymbol{R}}_j^\top \boldsymbol{A}_{ji} + \boldsymbol{A}_{ij}\widehat{\boldsymbol{R}}_j\widehat{\boldsymbol{R}}_i^\top\right).$$

We employ the following generalized projected power iteration scheme [34, 38, 63]:

$$\boldsymbol{R}_i^{(t+1)} = \mathcal{P}\left(\sum_{j=1}^{n}\boldsymbol{A}_{ij}\boldsymbol{R}_j^{(t)}\right) = \boldsymbol{U}_i^{(t)}\left(\boldsymbol{V}_i^{(t)}\right)^\top \tag{9}$$

where the initialization is chosen as $\boldsymbol{R}^{(0)} = \boldsymbol{Z}$, i.e., $\boldsymbol{R}_i^{(0)} = \boldsymbol{I}_d$. In practice, one can also use spectral method to produce an initialization which is close to the ground truth solution. Here $\boldsymbol{U}_i^{(t)}$ and $\boldsymbol{V}_i^{(t)}$ are left/right singular vectors of $\sum_{j=1}^n \boldsymbol{A}_{ij}\boldsymbol{R}_j^{(t)}$ respectively. More precisely, we have

$$\begin{aligned}\sum_{j=1}^{n}\boldsymbol{A}_{ij}\boldsymbol{R}_j^{(t)} &= \boldsymbol{U}_i^{(t)}\boldsymbol{\Sigma}_i^{(t)}\left(\boldsymbol{V}_i^{(t)}\right)^\top \\ &= \boldsymbol{U}_i^{(t)}\boldsymbol{\Sigma}_i^{(t)}\left(\boldsymbol{U}_i^{(t)}\right)^\top \cdot \boldsymbol{U}_i^{(t)}\left(\boldsymbol{V}_i^{(t)}\right)^\top \\ &= \boldsymbol{\Lambda}_{ii}^{(t+1)}\boldsymbol{R}_i^{(t+1)}\end{aligned}$$

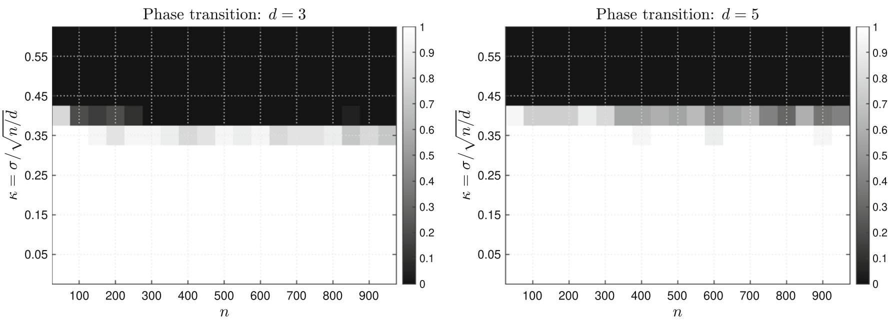

Fig. 2 The phase transition plot for the tightness of SDP relaxation when $d = 3$ and 5. White region: the tightness of the SDP holds successfully; black region: the tightness fails to hold. One can see if the noise is small (approximately $\sigma = \kappa\sqrt{n/d}$ with $\kappa \leq 0.35$), the SDP relaxation is tight

and $\boldsymbol{\Lambda}_{ii}^{(t)} = \boldsymbol{U}_i^{(t)}\boldsymbol{\Sigma}_i^{(t)}(\boldsymbol{U}_i^{(t)})^\top$ is a symmetric matrix. The fixed point $\boldsymbol{R}^{(\infty)}$ of this iteration satisfies

$$\sum_{j=1}^{n} \boldsymbol{A}_{ij}\boldsymbol{R}_j^{(\infty)} = \boldsymbol{\Lambda}_{ii}^{(\infty)}\boldsymbol{R}_i^{(\infty)}, \quad 1 \leq i \leq n$$

which is actually the first-order necessary condition as discussed in Lemma 1. If the fixed point is found, it remains to show the $(d+1)$-th smallest eigenvalue of $\boldsymbol{\Lambda}^\infty - \boldsymbol{A}$ is strictly positive, i.e, $\lambda_{d+1}(\boldsymbol{\Lambda}^\infty - \boldsymbol{A}) > 0$, in order to confirm $\boldsymbol{R}^{(\infty)}(\boldsymbol{R}^{(\infty)})^\top$ is the optimal solution to (SDP). In the experiment, we treat $\boldsymbol{R}^{(t)}(\boldsymbol{R}^{(t)})^\top$ as the global optimal solution to (SDP) if

$$\left\|(\boldsymbol{\Lambda}^{(t)} - \boldsymbol{A})\boldsymbol{R}^{(t)}\right\|_{\mathrm{op}} < 10^{-6}, \quad \lambda_{d+1}\left(\boldsymbol{\Lambda}^{(t)} - \boldsymbol{A}\right) > 10^{-8}. \tag{10}$$

The iteration stops if either iteration stabilizes:

$$\left\|\boldsymbol{R}^{(t+1)} - \boldsymbol{R}^{(t)}\right\|_F \leq 10^{-12}$$

or the number of iteration reaches 500. Typically, the iteration stops within 100 steps if the noise σ is small and does not converge within 500 steps for large noise, see Fig. 3 for a detailed discussion on how σ affects the iteration number.

In this experiment, we let $d = 3$ or 5 and $\sigma = \kappa\sqrt{n/d}$. The parameters (κ, n) are set to be $0 \leq \kappa \leq 0.6$ and $100 \leq n \leq 1000$. For each pair of (κ, n), we run 20 instances and calculate the proportion of successful cases, i.e., (10) holds. From Fig. 2, we see that if $\kappa < 0.35$, the SDP is tight, i.e., it recovers the global minimizer to the least squares objective function. The phase transition plot does not depend heavily on the parameter d. This confirms our conjecture numerically that if $\sigma \lesssim \sqrt{n/d}$ (modulo a log factor), the SDP relaxation is tight for the O(d) synchronization under additive Gaussian noise.

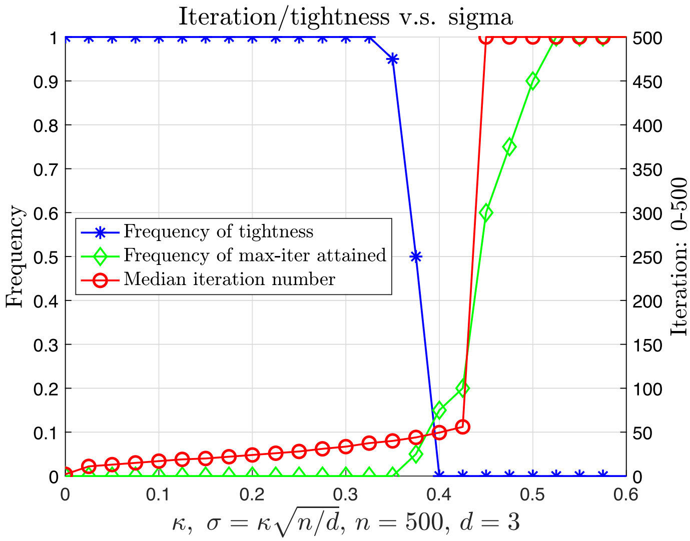

Fig. 3 The impact of σ on the iteration number and the global convergence of the generalized power method with $n = 500$, $d = 3$, and $p = d$. The noise level κ varies from 0 to 0.6 with stepsize 0.025. For each κ, the iteration number needed to achieve the stopping criterion (either $\|\boldsymbol{R}^{t+1} - \boldsymbol{R}^t\|_F \leq 10^{-12}$ or the number of iteration reaches 500) is recorded for 20 experiments. Once the iteration stops, we use duality theory to check if the recovered solution is globally optimal. The figure above indicates that all instances converge within 50 iterations for $\kappa \leq 0.35$ and achieve the globally optimal solutions; for $0.35 < \kappa \leq 0.425$, the algorithm usually stops within 100 iterations in most cases while only a small fraction of cases attain the maximum iteration number. But they fail to achieve the globally optimal solutions; for $\kappa \geq 0.45$, the algorithm neither stops within 500 iterations nor provides a globally optimal solution in most cases

4.2 Phase transition plot for nonconvex low-rank optimization

Instead of applying Riemannian gradient method to (BM), we employ projected power method to show how the convergence depends on the noise level. Here the generalized power method is viewed as conditional gradient method. We *randomly* initialize each $\boldsymbol{S}_i^{(0)}$ by creating a $d \times p$ Gaussian random matrix and extracting the random row space via QR decomposition. Then we perform

$$\boldsymbol{S}_i^{(t+1)} = \mathcal{P}\left(\sum_{j=1}^{n} \boldsymbol{A}_{ij}\boldsymbol{S}_j^{(t)}\right)$$

and the projection operator $\mathcal{P}$ is defined in (9).

After the iterates stabilize, we use (10) to verify the global optimality and tightness of the solution. Here we set $p = 2d$ and for each pair of κ and n, we still run 20 experiments and calculate the proportion of successful instances. Compared with Figs. 2, 4 provides highly similar phase transition plots for both $d = 3$ and $d = 5$. This is a strong indicator that the objective function is likely to have a benign landscape even if $\sigma \lesssim \sqrt{n/d}$, which is much more optimistic than our current theoretical bound.

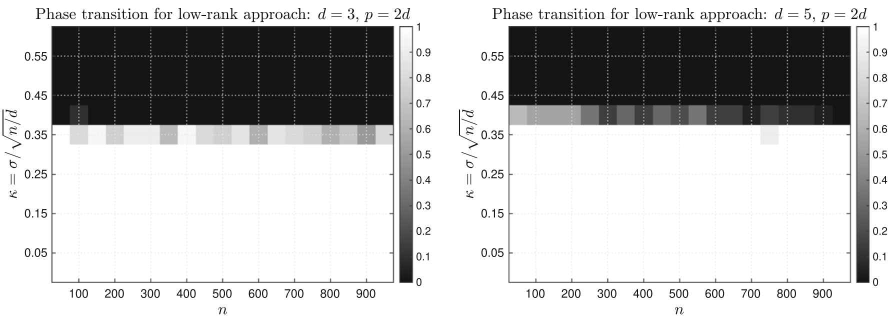

Fig. 4 The phase transition plot for the Burer-Monteiro approach to successfully find the global optimal solution when $d = 3$ and 5. White region: the BM approach finds the global maximizer successfully; black region: the BM approach fails. One can see if the noise is small (approximately $\sigma = \kappa\sqrt{n/d}$ with $\kappa \le 0.35$), the BM approach with random initialization produces the global optimal solution

4.3 Running time

One benefit of using either generalized power method or the Burer-Monteiro approach is the high computational efficiency. To provide such a comparison with the standard SDP solver, we consider the O(d) synchronization with additive Gaussian noise $\boldsymbol{A} = \boldsymbol{Z}\boldsymbol{Z}^\top + \sigma\boldsymbol{W}$ with $\sigma = \kappa\sqrt{n/d}$ and $\kappa = 0.2$. Then we apply the generalized power method ($p = d$) with random initialization and the benchmark SDP solver SDPNAL+ [51, 60] to solve the (SDP) relaxation exactly. The parameter n varies from 100 to 500, and for each n, we apply these two methods to 10 instances and record the computational time. We also compute the relative error between the solutions from the generalized power method (GPM) and SDPNAL+, and the relative error is of order 10^{-5} for all n in our experiment. The running time comparison is illustrated in Fig. 5. We can see that as the problem size n gets larger, the average running time for SDPNAL+ grows significantly faster than that for the GPM. For $d = 5$ and $n = 500$, it takes SDPNAL+ approximately 6 minutes on average to solve one instance exactly while it takes less than 1 second for the GPM. This demonstrates the GPM produces high-quality solutions to the SDP relaxation while also enjoying high computational efficiency.

5 Proofs

5.1 The roadmap of the proof

The proof consists of several sections and some parts are rather technical. Thus we provide a roadmap of the proof here. For both the analysis of convex relaxation and the Burer-Monteiro factorization, the key is to analyze the objective function $f(\boldsymbol{S})$ defined in (BM) for $\boldsymbol{S} \in \mathrm{St}(d, p)^{\otimes n}$ for $p > d$. Our analysis consists of several steps:

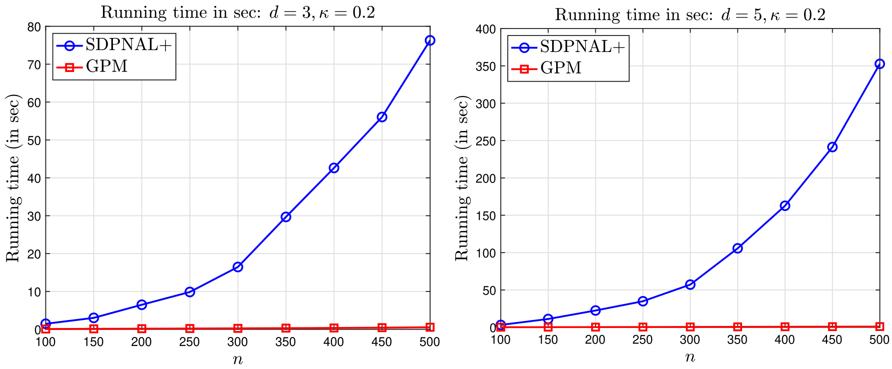

Fig. 5 Comparison on the running time between the SDPNAL+ [51] and the GPM for solving the orthogonal group synchronization with $d = 3$ and 5. The experiments are performed with MATLAB R2021a on MacBook Pro (16-inch 2019, 2.6 GHz 6-Core Intel Core i7). We use the default settings for the SDPNAL+ and in particular, the accuracy tolerance is 10^{-6} in the relative KKT residual [51]

(a) For (BM), we first provide a sufficient condition to certify the global optimality and tightness of a candidate solution $\boldsymbol{S}$ by using the duality theory in convex optimization. This is given in Proposition 1.
(b) We show that $\boldsymbol{S}$ is the global maximizer of (BM) if $\boldsymbol{S}$ is a second-order critical point and satisfies the proximity condition, i.e., $\boldsymbol{S}$ is sufficiently close to the *fully synchronized state*, i.e., $\boldsymbol{S}_i = \boldsymbol{S}_j, \forall i \neq j$. Moreover, the rank of $\boldsymbol{S}$ equals d and thus it is tight. This is summarized by Proposition 2.
(c) For the proof of the tightness of convex relaxation, we show that the *global maximizer* of (BM) with $p > d$ must satisfy the proximity condition, see Proposition 3; for the Burer-Monteiro approach with $p \geq 2d + 1$, we prove that *all* the second order critical points (SOCP) of (BM) satisfy the proximity condition, as shown in the Proposition 4.
(d) Combining all the supporting results together finishes the proof.

The idea of proof is mainly inspired by [7, 8] which focus on the $\mathbb{Z}_2$- and angular synchronization. However, due to the non-commutativity of $\mathrm{O}(d)$ for $d \geq 3$, several parts require quite different treatments. Now we present the first proposition which gives a sufficient condition to guarantee the global optimality and tightness, and establish the equivalence of the global maximizers among the three optimization programs (P), (SDP), and (BM). Without loss of generality, we assume $\boldsymbol{A} = \boldsymbol{Z}\boldsymbol{Z}^\top + \boldsymbol{\Delta}$, i.e., $\boldsymbol{A}_{ij} = \boldsymbol{I}_d + \boldsymbol{\Delta}_{ij}$, where $\boldsymbol{Z}^\top = [\boldsymbol{I}_d, \cdots, \boldsymbol{I}_d] \in \mathbb{R}^{d\times nd}$ from now on.

Proposition 1 *Let $\boldsymbol{\Lambda}$ be an $nd \times nd$ block diagonal matrix $\boldsymbol{\Lambda} = \mathrm{blkdiag}(\boldsymbol{\Lambda}_{11}, \cdots, \boldsymbol{\Lambda}_{nn})$. Suppose $\boldsymbol{\Lambda}$ satisfies*

$$(\boldsymbol{\Lambda} - \boldsymbol{A})\boldsymbol{S} = 0, \quad \boldsymbol{\Lambda} - \boldsymbol{A} \succeq 0, \tag{11}$$

for some $\boldsymbol{S} \in \mathrm{St}(d, p)^{\otimes n}$, then $\boldsymbol{S}\boldsymbol{S}^\top \in \mathbb{R}^{nd\times nd}$ is a global optimal solution to the SDP relaxation (SDP). Moreover, $\boldsymbol{S}$ is the unique global optimal solution to (P) if the following additional rank assumption holds

$$\operatorname{rank}(\boldsymbol{\Lambda} - \boldsymbol{A}) = (n-1)d. \tag{12}$$

The condition (11) provides a sufficient condition for $\boldsymbol{X} = \boldsymbol{S}\boldsymbol{S}^\top$ to be one global maximizer to (SDP) and (BM). The condition (12) characterizes when the solution $\boldsymbol{S} \in \mathbb{R}^{p\times nd}$ is of rank d and unique. In particular, if $\operatorname{rank}(\boldsymbol{S}) = d$, then $\boldsymbol{S}$ is actually the global maximizer to (P).

The next step is to show all the second order critical points, i.e., those points whose Riemannian gradient equals 0 and Hessian is positive semidefinite, are actually global maximizers if they are close to the fully synchronized state. It suffices to show that those SOCPs satisfy the global optimality condition (11) and (12). In fact, if $\boldsymbol{S}$ is a first order critical point, we immediately have $(\boldsymbol{\Lambda} - \boldsymbol{A})\boldsymbol{S} = 0$ for some block-diagonal matrix $\boldsymbol{\Lambda} \in \mathbb{R}^{nd\times nd}$.

Lemma 1 (First order critical point) *The first order critical point of $f(\boldsymbol{S})$ satisfies:*

$$\sum_{j=1}^n \boldsymbol{A}_{ij}\boldsymbol{S}_j = \boldsymbol{\Lambda}_{ii}\boldsymbol{S}_i, \quad \boldsymbol{S}_i \in \operatorname{St}(d,p)$$

where $\boldsymbol{\Lambda}_{ii}$ equals

$$\boldsymbol{\Lambda}_{ii} := \frac{1}{2}\sum_{j=1}^n \left(\boldsymbol{S}_i\boldsymbol{S}_j^\top \boldsymbol{A}_{ji} + \boldsymbol{A}_{ij}\boldsymbol{S}_j\boldsymbol{S}_i^\top \right). \tag{13}$$

The proof of Lemma 1 is given in Sect. 5.2. Lemma 1 shows that $\boldsymbol{\Lambda}_{ii}$ depends on $\boldsymbol{S}$ and is completely determined by $(\boldsymbol{\Lambda} - \boldsymbol{A})\boldsymbol{S} = 0$. As a result, it suffices to prove that $\boldsymbol{\Lambda} - \boldsymbol{A} \succeq 0$ for some second order critical points which obey the *proximity condition*, i.e., $\boldsymbol{S}$ is sufficiently close to $\boldsymbol{Z}$. To quantify this closeness, we introduce the following distance: given any $\boldsymbol{S} \in \operatorname{St}(d,p)^{\otimes n}$, the distance between $\boldsymbol{S}$ and the fully synchronized state is defined by

$$\begin{aligned} d_F(\boldsymbol{S},\boldsymbol{Z}) &:= \min_{\boldsymbol{Q}\in\mathbb{R}^{d\times p}, \boldsymbol{Q}\boldsymbol{Q}^\top = \boldsymbol{I}_d} \|\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q}\|_F \\ &= \min_{\boldsymbol{Q}\in\mathbb{R}^{d\times p}, \boldsymbol{Q}\boldsymbol{Q}^\top = \boldsymbol{I}_d} \sqrt{\sum_{i=1}^n \|\boldsymbol{S}_i - \boldsymbol{Q}\|_F^2} \end{aligned} \tag{14}$$

where $(\boldsymbol{Z}\boldsymbol{Q})^\top = [\boldsymbol{Q}^\top, \cdots, \boldsymbol{Q}^\top] \in \mathbb{R}^{p\times nd}$. For the rest of the paper, we will let $\boldsymbol{Q}$ be the $d \times p$ partial orthogonal matrix which minimizes (14). In fact, the minimizer equals $\boldsymbol{Q} = \mathcal{P}(\boldsymbol{Z}^\top\boldsymbol{S})$ where $\boldsymbol{Z}^\top\boldsymbol{S} = \sum_{j=1}^n \boldsymbol{S}_j \in \mathbb{R}^{d\times p}$ and $\mathcal{P}(\cdot)$ is defined in (9).

Condition 5 (Proximity condition) *A feasible solution $\boldsymbol{S} \in \operatorname{St}(d,p)^{\otimes n}$ satisfies the proximity condition if*

$$d_F(\boldsymbol{S},\boldsymbol{Z}) \le \delta\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\text{op}} \tag{15}$$

for some constant $\delta > 0$.

The next Proposition is the core of the whole proof, stating that any SOCPs satisfying the proximity condition (15) are global maximizers to (P) and (SDP).

Proposition 2 (bf Global optimality of (P) and (SDP)) *For a second order critical point* $\boldsymbol{S}$ *satisfying (15), it is the unique global maximizer to both (P) and (SDP) if*

$$n \geq \frac{3\delta^2 d\|\boldsymbol{\Delta}\|_{\mathrm{op}}^2}{2n} + \max_{1\leq i\leq n}\left\|\sum_{j\neq i}\boldsymbol{\Delta}_{ij}\boldsymbol{S}_j\right\|_{\mathrm{op}} + \|\boldsymbol{\Delta}\|_{\mathrm{op}} \tag{16}$$

where

$$\max_{1\leq i\leq n}\left\|\sum_{j\neq i}\boldsymbol{\Delta}_{ij}\boldsymbol{S}_j\right\|_{\mathrm{op}} \leq \delta\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}}\max_{1\leq i\leq n}\|\boldsymbol{\Delta}_i\|_{\mathrm{op}} + \max_{1\leq i\leq n}\left\|\boldsymbol{\Delta}_i^\top\boldsymbol{Z}\right\|_{\mathrm{op}}. \tag{17}$$

In other words, the global optimality of $\boldsymbol{S}$ *is guaranteed by*

$$n \geq \frac{3\delta^2 d\|\boldsymbol{\Delta}\|_{\mathrm{op}}^2}{2n} + \delta\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}}\max_{1\leq i\leq n}\|\boldsymbol{\Delta}_i\|_{\mathrm{op}} + \max_{1\leq i\leq n}\left\|\boldsymbol{\Delta}_i^\top\boldsymbol{Z}\right\|_{\mathrm{op}} + \|\boldsymbol{\Delta}\|_{\mathrm{op}}.$$

Remark 1 Proposition 2 provides a simple criterion to verify a near-fully synchronized state is the global optimal solution. However, the estimation of $\max_{1\leq i\leq n}\left\|\sum_{j\neq i}\boldsymbol{\Delta}_{ij}\boldsymbol{S}_j\right\|_{\mathrm{op}}$ is not tight which leads to the suboptimal bound in the main theorems. The major difficulty results from the complicated statistical dependence between $\boldsymbol{\Delta}$ and any second-order critical points $\boldsymbol{S}$. This is well worth further investigation for $\mathrm{O}(d)$.

Now we present two propositions which demonstrate that any global maximizers and second-order critical points to (BM) satisfy (15) for some $\delta > 0$.

(i) *Convex relaxation* For the tightness of SDP relaxation, we show that the global maximizer to (P) must satisfy (15) with $\delta = 4$.

Proposition 3 (Proximity condition for convex relaxation) *The global maximizers to (BM) satisfy*

$$d_F(\boldsymbol{S},\boldsymbol{Z}) \leq \delta\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}}, \quad \delta = 4.$$

This proposition essentially ensures that any global maximizer to (BM) is close to the fully synchronized state and its distance depends on the noise strength.

(ii) *Low-rank approach* For the Burer-Monteiro approach, we prove that if $p \geq 2d+1$, all the local maximizers of (BM) satisfy (15) with δ which depends on p, d, and γ.

Proposition 4 (Proximity condition for low-rank approach) *Suppose $p \geq 2d+1$. All the second-order critical points $\boldsymbol{S}$ of $f(\boldsymbol{S})$ in (BM) satisfy*

$$d_F(\boldsymbol{S}, \boldsymbol{Z}) \leq \delta\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}}, \quad \delta = \frac{\left(2+\sqrt{5}\right)(p+d)\gamma}{p-2d}$$

where

$$\gamma := \frac{\|\operatorname{Tr}_d(\boldsymbol{\Delta})\|_{\mathrm{op}}}{\|\boldsymbol{\Delta}\|_{\mathrm{op}}} \vee 1$$

and $\operatorname{Tr}_d(\boldsymbol{\Delta}) = [\operatorname{Tr}(\boldsymbol{\Delta}_{ij})]_{ij} \in \mathbb{R}^{n\times n}$ denotes the partial trace of $\boldsymbol{\Delta}$.

Remark 2 If $\boldsymbol{\Delta}$ is a symmetric Gaussian random matrix, then $\operatorname{Tr}_d(\boldsymbol{\Delta})$ is an $n \times n$ Gaussian random matrix whose entry is $\mathcal{N}(0, d)$ and $\|\operatorname{Tr}_d(\boldsymbol{\Delta})\|_{\mathrm{op}} = (1+o(1))\|\boldsymbol{\Delta}\|_{\mathrm{op}}$ holds.

We defer the proof of Proposition 1, 2, 3 and 4 to Sects. 5.3, 5.4, 5.5 and 5.6 respectively. Now we provide a proof of Theorems 1 and 3 by using the aforementioned propositions.

Proof of Theorem 1 To prove the tightness of convex relaxation, we first consider the global maximizer to (BM) which is also a second-order critical point. By Proposition 3, we have $d_F(\boldsymbol{S}, \boldsymbol{Z}) \leq \delta\sqrt{n^{-1}d}\|\boldsymbol{\Delta}\|_{\mathrm{op}}$ with $\delta = 4$. With Proposition 2, we immediately have Theorem 1. □

Proof of Theorem 3 To analyze the landscape of (BM), we invoke Proposition 4 which states that *all* the second-order critical points (SOCP) are essentially close to the fully synchronized state. Now it suffices to show that all SOCPs are global maximizers to (SDP) and (P) and the global maximizer is unique under the assumption of Theorem 3. This is fortunately guaranteed by Proposition 2. □

5.2 Riemannian gradient and Hessian matrix

We start with analyzing the SOCPs of $f(\boldsymbol{S})$ by first computing its Riemannian gradient and Hessian. The calculation involves the tangent space at $\boldsymbol{S}_i \in \operatorname{St}(d, p)$ which is given by

$$T_{\boldsymbol{S}_i}(\mathcal{M}) := \left\{\boldsymbol{Y}_i \in \mathbb{R}^{d\times p} : \boldsymbol{S}_i\boldsymbol{Y}_i^\top + \boldsymbol{Y}_i\boldsymbol{S}_i^\top = 0\right\}, \quad \mathcal{M} := \operatorname{St}(d, p). \tag{18}$$

In other words, $\boldsymbol{S}_i\boldsymbol{Y}_i^\top$ is an anti-symmetric matrix if $\boldsymbol{Y}_i \in \mathbb{R}^{d\times p}$ is an element in the tangent space.

Proof of Lemma 1 Recall the objective function $f(\boldsymbol{S}) = \sum_{i=1}^n\sum_{j=1}^n\langle \boldsymbol{S}_i, \boldsymbol{A}_{ij}\boldsymbol{S}_j\rangle$ in (BM) where $\boldsymbol{S}_i \in \operatorname{St}(d, p)$. We take the gradient w.r.t. $\boldsymbol{S}_i$ in the Euclidean space.

$$\frac{\partial f}{\partial \boldsymbol{S}_i} = \sum_{j\neq i}\boldsymbol{A}_{ij}\boldsymbol{S}_j.$$

The Riemannian gradient w.r.t. $\boldsymbol{S}_i$ is

$$\nabla_{\boldsymbol{S}_i} f = \mathrm{Proj}_{T_{\boldsymbol{S}_i}(\mathcal{M})}\left(\sum_{j\neq i}\boldsymbol{A}_{ij}\boldsymbol{S}_j\right) = \sum_{j\neq i}\boldsymbol{A}_{ij}\boldsymbol{S}_j - \frac{1}{2}\sum_{j\neq i}\left(\boldsymbol{S}_i\boldsymbol{S}_j^\top\boldsymbol{A}_{ji} + \boldsymbol{A}_{ij}\boldsymbol{S}_j\boldsymbol{S}_i^\top\right)\boldsymbol{S}_i \tag{19}$$

by projecting $\frac{\partial f}{\partial \boldsymbol{S}_i}$ onto the tangent space $T_{\boldsymbol{S}_i}(\mathcal{M})$ at $\boldsymbol{S}$, as shown in [3, Equation (3.35)]:

$$\mathrm{Proj}_{T_{\boldsymbol{S}_i}(\mathcal{M})}(\boldsymbol{\Pi}) = \boldsymbol{\Pi} - \frac{1}{2}\left(\boldsymbol{\Pi}\boldsymbol{S}_i^\top + \boldsymbol{S}_i\boldsymbol{\Pi}^\top\right)\boldsymbol{S}_i$$

where $\boldsymbol{\Pi} \in \mathbb{R}^{d\times p}$ and the matrix manifold $\mathcal{M}$ is $\mathrm{St}(d, p)$. Setting $\nabla_{\boldsymbol{S}_i} f = 0$ and multiplying both sides by $\boldsymbol{S}_i^\top$ gives

$$\sum_{j\neq i}\boldsymbol{A}_{ij}\boldsymbol{S}_j\boldsymbol{S}_i^\top = \sum_{j\neq i}\boldsymbol{S}_i\boldsymbol{S}_j^\top\boldsymbol{A}_{ji}$$

since $\boldsymbol{S}_i\boldsymbol{S}_i^\top = \boldsymbol{I}_d$. Therefore, it holds that

$$\sum_{j\neq i}\boldsymbol{A}_{ij}\boldsymbol{S}_j = \sum_{j\neq i}\boldsymbol{A}_{ij}\boldsymbol{S}_j\boldsymbol{S}_i^\top\boldsymbol{S}_i \Longrightarrow \sum_{j=1}^{n}\boldsymbol{A}_{ij}\boldsymbol{S}_j = \sum_{j=1}^{n}\boldsymbol{A}_{ij}\boldsymbol{S}_j\boldsymbol{S}_i^\top\boldsymbol{S}_i.$$

By the definition of $\boldsymbol{\Lambda}_{ii}$ in (13), we have

$$\sum_{j=1}^{n}\boldsymbol{A}_{ij}\boldsymbol{S}_j = \boldsymbol{\Lambda}_{ii}\boldsymbol{S}_i,$$

In other words, $(\boldsymbol{\Lambda} - \boldsymbol{A})\boldsymbol{S} = 0$ where $\boldsymbol{\Lambda} = \mathrm{blkdiag}(\boldsymbol{\Lambda}_{11}, \cdots, \boldsymbol{\Lambda}_{nn})$. □

Next, we compute the Riemannian Hessian and prove that $\boldsymbol{\Lambda} \succeq 0$ for any second order critical point.

Lemma 2 *The quadratic form associated to the Hessian matrix of (BM) is*

$$\dot{\boldsymbol{S}} : \nabla^2_{\partial\boldsymbol{S}\partial\boldsymbol{S}} f(\boldsymbol{S}) : \dot{\boldsymbol{S}} = -\sum_{i=1}^{n}\left\langle\boldsymbol{\Lambda}_{ii}, \dot{\boldsymbol{S}}_i\dot{\boldsymbol{S}}_i^\top\right\rangle + \sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle\boldsymbol{A}_{ij}, \dot{\boldsymbol{S}}_i\dot{\boldsymbol{S}}_j^\top\right\rangle$$

where $\dot{\boldsymbol{S}}^\top = [\dot{\boldsymbol{S}}_1^\top, \cdots, \dot{\boldsymbol{S}}_n^\top] \in \mathbb{R}^{p\times nd}$ and $\dot{\boldsymbol{S}}_i \in \mathbb{R}^{d\times p}$ is an element on the tangent space of Stiefel manifold at $\boldsymbol{S}_i$. In other words, if $\boldsymbol{S}$ is a second order critical point, it must satisfy:

$$\sum_{i=1}^{n}\left\langle\boldsymbol{\Lambda}_{ii}, \dot{\boldsymbol{S}}_i\dot{\boldsymbol{S}}_i^\top\right\rangle \geq \sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle\boldsymbol{A}_{ij}, \dot{\boldsymbol{S}}_i\dot{\boldsymbol{S}}_j^\top\right\rangle. \tag{20}$$

Proof Recall the Riemannian gradient w.r.t. $\boldsymbol{S}_i$ is given by

$$\nabla_{\boldsymbol{S}_i} f = \sum_{j\neq i} \boldsymbol{A}_{ij}\boldsymbol{S}_j - \frac{1}{2}\left(\sum_{j\neq i} \boldsymbol{S}_i\boldsymbol{S}_j^\top \boldsymbol{A}_{ji} + \boldsymbol{A}_{ij}\boldsymbol{S}_j\boldsymbol{S}_i^\top\right)\boldsymbol{S}_i.$$

Let $\dot{\boldsymbol{S}}_i$ be a matrix on the tangent space at $\boldsymbol{S}_i$:

$$\lim_{t\to 0} \frac{\nabla_{\boldsymbol{S}_i} f(\boldsymbol{S}+t\dot{\boldsymbol{S}}_i) - \nabla_{\boldsymbol{S}_i} f(\boldsymbol{S})}{t} = -(\boldsymbol{A}_{ii} - \boldsymbol{I}_d)\dot{\boldsymbol{S}}_i - \frac{1}{2}\sum_{j\neq i}\left(\dot{\boldsymbol{S}}_i\boldsymbol{S}_j^\top \boldsymbol{A}_{ji} + \boldsymbol{A}_{ij}\boldsymbol{S}_j\dot{\boldsymbol{S}}_i^\top\right)\boldsymbol{S}_i$$

where $(\boldsymbol{S}+t\dot{\boldsymbol{S}}_i)^\top = [\boldsymbol{S}_1^\top, \cdots, (\boldsymbol{S}_i + t\dot{\boldsymbol{S}}_i)^\top, \cdots, \boldsymbol{S}_n^\top]$ and $\boldsymbol{S}_i\boldsymbol{S}_i^\top = \boldsymbol{I}_d$. As a result, the quadratic form associated to the Riemannian Hessian is

$$\begin{aligned}
-\dot{\boldsymbol{S}}_i : \nabla^2_{\partial \boldsymbol{S}_i \partial \boldsymbol{S}_i} f(\boldsymbol{S}) : \dot{\boldsymbol{S}}_i &= \left\langle (\boldsymbol{A}_{ii} - \boldsymbol{I}_d)\dot{\boldsymbol{S}}_i + \frac{1}{2}\sum_{j\neq i}\left(\dot{\boldsymbol{S}}_i\boldsymbol{S}_j^\top \boldsymbol{A}_{ji} + \boldsymbol{A}_{ij}\boldsymbol{S}_j\dot{\boldsymbol{S}}_i^\top\right)\boldsymbol{S}_i, \dot{\boldsymbol{S}}_i\right\rangle \\
&= \left\langle \boldsymbol{A}_{ii} - \boldsymbol{I}_d, \dot{\boldsymbol{S}}_i\dot{\boldsymbol{S}}_i^\top\right\rangle \\
&\quad + \frac{1}{2}\sum_{j\neq i}\left\langle \dot{\boldsymbol{S}}_i\boldsymbol{S}_j^\top\boldsymbol{A}_{ji}, \dot{\boldsymbol{S}}_i\boldsymbol{S}_i^\top\right\rangle + \frac{1}{2}\sum_{j\neq i}\left\langle \boldsymbol{A}_{ij}\boldsymbol{S}_j\dot{\boldsymbol{S}}_i^\top, \dot{\boldsymbol{S}}_i\boldsymbol{S}_i^\top\right\rangle \\
&= \left\langle \boldsymbol{A}_{ii} - \boldsymbol{I}_d, \dot{\boldsymbol{S}}_i\dot{\boldsymbol{S}}_i^\top\right\rangle + \frac{1}{2}\sum_{j\neq i}\left\langle \boldsymbol{A}_{ij}\boldsymbol{S}_j\dot{\boldsymbol{S}}_i^\top, \boldsymbol{S}_i\dot{\boldsymbol{S}}_i^\top + \dot{\boldsymbol{S}}_i\boldsymbol{S}_i^\top\right\rangle \\
&= \langle \boldsymbol{A}_{ii} - \boldsymbol{I}_d, \dot{\boldsymbol{S}}_i\dot{\boldsymbol{S}}_i^\top\rangle
\end{aligned}$$

where $\boldsymbol{S}_i\dot{\boldsymbol{S}}_i^\top + \dot{\boldsymbol{S}}_i\boldsymbol{S}_i^\top = 0$ since $\dot{\boldsymbol{S}}_i$ is on $T_{\boldsymbol{S}_i}(\mathcal{M})$.

For the mixed partial derivative, we have

$$\lim_{t\to 0} \frac{\nabla_{\boldsymbol{S}_i} f(\boldsymbol{S}+t\dot{\boldsymbol{S}}_j) - \nabla_{\boldsymbol{S}_i} f(\boldsymbol{S})}{t} = \boldsymbol{A}_{ij}\dot{\boldsymbol{S}}_j - \frac{1}{2}\left(\boldsymbol{S}_i\dot{\boldsymbol{S}}_j^\top \boldsymbol{A}_{ji} + \boldsymbol{A}_{ij}\dot{\boldsymbol{S}}_j\boldsymbol{S}_i^\top\right)\boldsymbol{S}_i$$

for $j \neq i$. Thus

$$\begin{aligned}
\dot{\boldsymbol{S}}_i : \nabla^2_{\partial \boldsymbol{S}_i \partial \boldsymbol{S}_j} f(\boldsymbol{S}) : \dot{\boldsymbol{S}}_j &= \left\langle \boldsymbol{A}_{ij}\dot{\boldsymbol{S}}_j - \frac{1}{2}\left(\boldsymbol{S}_i\dot{\boldsymbol{S}}_j^\top \boldsymbol{A}_{ji} + \boldsymbol{A}_{ij}\dot{\boldsymbol{S}}_j\boldsymbol{S}_i^\top\right)\boldsymbol{S}_i, \dot{\boldsymbol{S}}_i\right\rangle \\
&= \langle \boldsymbol{A}_{ij}\dot{\boldsymbol{S}}_j, \dot{\boldsymbol{S}}_i\rangle - \frac{1}{2}\left\langle \boldsymbol{S}_i\dot{\boldsymbol{S}}_j^\top\boldsymbol{A}_{ji} + \boldsymbol{A}_{ij}\dot{\boldsymbol{S}}_j\boldsymbol{S}_i^\top, \dot{\boldsymbol{S}}_i\boldsymbol{S}_i^\top\right\rangle \\
&= \langle \boldsymbol{A}_{ij}\dot{\boldsymbol{S}}_j, \dot{\boldsymbol{S}}_i\rangle - \frac{1}{2}\left\langle \boldsymbol{A}_{ij}\dot{\boldsymbol{S}}_j\boldsymbol{S}_i^\top, \boldsymbol{S}_i\dot{\boldsymbol{S}}_i^\top + \dot{\boldsymbol{S}}_i\boldsymbol{S}_i^\top\right\rangle \\
&= \langle \boldsymbol{A}_{ij}\dot{\boldsymbol{S}}_j, \dot{\boldsymbol{S}}_i\rangle.
\end{aligned}$$

Taking the sum of $\dot{\boldsymbol{S}}_i : \nabla^2_{\partial \boldsymbol{S}_i \partial \boldsymbol{S}_j} f(\boldsymbol{S}) : \dot{\boldsymbol{S}}_j$ over (i, j) gives

$$\dot{\boldsymbol{S}} : \nabla^2_{\partial \boldsymbol{S} \partial \boldsymbol{S}} f : \dot{\boldsymbol{S}} = -\sum_{i=1}^{n} \left\langle \boldsymbol{\Lambda}_{ii}, \dot{\boldsymbol{S}}_i \dot{\boldsymbol{S}}_i^\top \right\rangle + \sum_{i=1}^{n} \sum_{j=1}^{n} \left\langle \boldsymbol{A}_{ij}, \dot{\boldsymbol{S}}_i \dot{\boldsymbol{S}}_j^\top \right\rangle.$$

If $\boldsymbol{S}$ is a local maximizer of (BM), then $\dot{\boldsymbol{S}} : \nabla^2_{\partial \boldsymbol{S} \partial \boldsymbol{S}} f : \dot{\boldsymbol{S}} \leq 0$ holds for any $\dot{\boldsymbol{S}} \in (T_{\boldsymbol{S}_i}(\mathcal{M}))^{\otimes n}$. ◻

Suppose $\boldsymbol{S}$ is a local maximizer of (BM), then (20) implies that

$$\left\langle \boldsymbol{\Lambda}_{ii} - \boldsymbol{I}_d, \dot{\boldsymbol{S}}_i \dot{\boldsymbol{S}}_i^\top \right\rangle \geq 0, \quad \dot{\boldsymbol{S}}_i \in T_{\boldsymbol{S}_i}(\mathcal{M}).$$

Does it imply that $\boldsymbol{\Lambda}_{ii} \succeq \boldsymbol{I}_d$? The answer is yes if $p > d$. However, this is not longer true if $p = d$. For $p = d$, we are only able to prove that the sum of the smallest two eigenvalues is nonnegative.

Lemma 3 *Under $p > d$, suppose $\boldsymbol{S}$ is a local maximizer, then it holds*

$$\boldsymbol{\Lambda}_{ii} \succeq \boldsymbol{I}_d, \quad 1 \leq i \leq n.$$

Proof Note that $\boldsymbol{S}_i \in \mathbb{R}^{d \times p}$ with $p > d$ is a “fat" matrix. It means we can always find $\boldsymbol{v}_i \in \mathbb{R}^p$ which is perpendicular to all rows of $\boldsymbol{S}_i$, i.e., $\boldsymbol{S}_i \boldsymbol{v}_i = 0$. Without loss of generality, we assume $\boldsymbol{v}_i$ is a unit vector. Now we construct $\dot{\boldsymbol{S}}_i$ in the following form:

$$\dot{\boldsymbol{S}}_i = \boldsymbol{u}_i \boldsymbol{v}_i^\top$$

where $\boldsymbol{u}_i$ is an arbitrary vector in $\mathbb{R}^d$. It is easy to verify that

$$\boldsymbol{S}_i \dot{\boldsymbol{S}}_i^\top = \boldsymbol{S}_i \boldsymbol{v}_i \boldsymbol{u}_i^\top = 0,$$

which means $\dot{\boldsymbol{S}}_i$ is indeed an element in the tangent space of $\mathrm{St}(d, p)$ at $\boldsymbol{S}_i$.

Now, we have

$$\left\langle \boldsymbol{\Lambda}_{ii} - \boldsymbol{I}_d, \boldsymbol{u}_i \boldsymbol{u}_i^\top \right\rangle = \left\langle \boldsymbol{\Lambda}_{ii} - \boldsymbol{I}_d, \boldsymbol{u}_i \boldsymbol{v}_i^\top \boldsymbol{v}_i \boldsymbol{u}_i^\top \right\rangle = \left\langle \boldsymbol{\Lambda}_{ii} - \boldsymbol{I}_d, \dot{\boldsymbol{S}}_i \dot{\boldsymbol{S}}_i^\top \right\rangle \geq 0, \quad \forall \boldsymbol{u}_i \in \mathbb{R}^d$$

which implies that $\boldsymbol{\Lambda}_{ii} - \boldsymbol{I}_d \succeq 0$ and $\boldsymbol{\Lambda}_{ii} \succeq \boldsymbol{I}_d$. ◻

5.3 Certifying global optimality via dual certificate

To guarantee the global optimality of a feasible solution, we will employ the standard tools from the literature in compressive sensing and low-rank matrix recovery. The core part is to construct the dual certificate which confirms that the proposed feasible solution and the dual certificate yield strong duality.

Proof of Proposition 1 We start from the convex optimization (SDP) and derive its dual program. First introduce the symmetric matrix $\boldsymbol{\Pi}_{ii} \in \mathbb{R}^{d\times d}$ as the dual variable corresponding to the constraint $\boldsymbol{X}_{ii} = \boldsymbol{I}_d$ and then get the Lagrangian function. Here we switch from maximization to minimization in (SDP) by changing the sign in the objective function.

$$\begin{aligned}\mathcal{L}(\boldsymbol{X}, \boldsymbol{\Pi}) &= \sum_{i=1}^{n} \langle \boldsymbol{\Pi}_{ii}, \boldsymbol{X}_{ii} - \boldsymbol{I}_d \rangle - \langle \boldsymbol{A}, \boldsymbol{X} \rangle \\ &= \langle \boldsymbol{\Pi} - \boldsymbol{A}, \boldsymbol{X} \rangle - \operatorname{Tr}(\boldsymbol{\Pi})\end{aligned}$$

where $\boldsymbol{\Pi} = \operatorname{blkdiag}(\boldsymbol{\Pi}_{11}, \cdots, \boldsymbol{\Pi}_{nn}) \in \mathbb{R}^{nd\times nd}$ and $\boldsymbol{X} \succeq 0$. If $\boldsymbol{\Pi} - \boldsymbol{A}$ is not positive semidefinite, taking the infimum w.r.t. $\boldsymbol{X} \succeq 0$ for the Lagrangian function gives negative infinity. Thus we require $\boldsymbol{\Pi} - \boldsymbol{A} \succeq 0$:

$$\inf_{\boldsymbol{X}\succeq 0} \mathcal{L}(\boldsymbol{X}, \boldsymbol{\Pi}) = -\operatorname{Tr}(\boldsymbol{\Pi}).$$

As a result, the dual program of (SDP) is equivalent to

$$\min_{\boldsymbol{\Pi}\in\mathbb{R}^{nd\times nd}} \operatorname{Tr}(\boldsymbol{\Pi}) \quad \text{such that} \quad \boldsymbol{\Pi} - \boldsymbol{A} \succeq 0, \quad \boldsymbol{\Pi} \text{ is block-diagonal.}$$

Weak duality in convex optimization [10] implies that $\operatorname{Tr}(\boldsymbol{\Pi}) \geq \langle \boldsymbol{A}, \boldsymbol{X} \rangle$. Moreover, $(\boldsymbol{X}, \boldsymbol{\Pi})$ is a primal-dual optimal solution (not necessarily unique) if the complementary slackness holds

$$\langle \boldsymbol{\Pi} - \boldsymbol{A}, \boldsymbol{X} \rangle = 0, \quad \boldsymbol{\Pi} - \boldsymbol{A} \succeq 0 \tag{21}$$

since (21) implies strong duality, i.e., $\operatorname{Tr}(\boldsymbol{\Pi}) = \langle \boldsymbol{A}, \boldsymbol{X} \rangle$ since $\boldsymbol{X}_{ii} = \boldsymbol{I}_d$. In fact, this condition (21) is equivalent to

$$(\boldsymbol{\Pi} - \boldsymbol{A})\boldsymbol{X} = 0, \quad \boldsymbol{\Pi} - \boldsymbol{A} \succeq 0 \tag{22}$$

because both $\boldsymbol{\Pi} - \boldsymbol{A}$ and $\boldsymbol{X}$ are positive semidefinite.

Let $\boldsymbol{S} \in \mathbb{R}^{nd\times p}$ be a feasible solution. Suppose there exists an $nd \times nd$ block diagonal matrix $\boldsymbol{\Lambda}$ and satisfies (11). The global optimality of $\boldsymbol{X} = \boldsymbol{S}\boldsymbol{S}^{\top}$ follows directly from (11) and (22). In addition, if the rank of $\boldsymbol{\Pi} - \boldsymbol{A}$ is $(n-1)d$, then the global optimizer to (SDP) is exactly rank-d. This is due to $(\boldsymbol{\Pi} - \boldsymbol{A})\boldsymbol{X} = 0$, implying that the rank of $\boldsymbol{X}$ is at most d but $\boldsymbol{X}_{ii} = \boldsymbol{I}_d$ guarantees $\operatorname{rank}(\boldsymbol{X}) \geq d$. This results in the tightness of (SDP) since the global optimal solution to the SDP is exactly rank-d and thus must be the global optimal solution to (P) as well.

Now we prove that if $\operatorname{rank}(\boldsymbol{\Pi} - \boldsymbol{A}) = (n-1)d$, then $\boldsymbol{X}$ is the *unique* maximizer. Let's prove it by contradiction. If not, then there exists another global maximizer $\widetilde{\boldsymbol{X}}$ such that $\langle \boldsymbol{A}, \widetilde{\boldsymbol{X}} \rangle = \operatorname{Tr}(\boldsymbol{\Pi}) = \langle \boldsymbol{\Pi}, \widetilde{\boldsymbol{X}} \rangle$ since the feasible solution $\boldsymbol{X}$ and $\widetilde{\boldsymbol{X}}$ achieve the same primal value due to the linearity of the objective function:

$$\langle \boldsymbol{\Pi} - \boldsymbol{A}, \widetilde{\boldsymbol{X}} \rangle = 0 \Longrightarrow (\boldsymbol{\Pi} - \boldsymbol{A})\widetilde{\boldsymbol{X}} = 0.$$

Since $\operatorname{rank}(\boldsymbol{\Pi} - \boldsymbol{A}) = (n-1)d$, thus $\operatorname{rank}(\widetilde{\boldsymbol{X}}) \le d$. Note that each diagonal block is $\boldsymbol{I}_d$ and it implies $\operatorname{rank}(\widetilde{\boldsymbol{X}}) = d$. This proves that $\widetilde{\boldsymbol{X}} = \boldsymbol{X}$ holds (modulo a global rotation in the column space) since $\boldsymbol{X}$ and $\widetilde{\boldsymbol{X}}$ are determined uniquely by the null space of $\boldsymbol{\Pi} - \boldsymbol{A}$. □

Proposition 1 indicates that in order to show that a first-order critical point of $f(\boldsymbol{S})$ is the unique global maximizer to (P), it suffices to guarantee $\boldsymbol{\Lambda} - \boldsymbol{A} \succeq 0$ and $\lambda_{d+1}(\boldsymbol{\Lambda} - \boldsymbol{A}) > 0$. This is equivalent to $\operatorname{rank}(\boldsymbol{\Lambda} - \boldsymbol{A}) = (n-1)d$ here since the first order necessary condition implies $(\boldsymbol{\Lambda} - \boldsymbol{A})\boldsymbol{S} = 0$ for $\boldsymbol{\Lambda}$ defined in (13) which means at least d eigenvalues of $\boldsymbol{\Lambda} - \boldsymbol{A}$ are zero. Now we can see that the key is to ensure $\boldsymbol{\Lambda} - \boldsymbol{A} \succeq 0$ for some first-order critical point $\boldsymbol{S}$ (i.e., those critical points which satisfy the proximity condition). Define the certificate matrix

$$\boldsymbol{C} := \boldsymbol{\Lambda} - \boldsymbol{A}, \quad \boldsymbol{C}_{ij} = \begin{cases} \boldsymbol{\Lambda}_{ii} - \boldsymbol{A}_{ii}, & i = j, \\ -\boldsymbol{A}_{ij}, & i \neq j, \end{cases} \tag{23}$$

for any given $\boldsymbol{S}$. From the definition, we know that any first-order critical points satisfy $\boldsymbol{C}\boldsymbol{S} = 0$.

5.4 Proof of Proposition 2

Lemma 4 *Suppose the proximity condition (15) holds, we have*

$$n \ge \sigma_{\max}\left(\boldsymbol{Z}^\top \boldsymbol{S}\right) \ge \sigma_{\min}\left(\boldsymbol{Z}^\top \boldsymbol{S}\right) \ge n - \frac{\delta^2 d \|\boldsymbol{\Delta}\|_{\mathrm{op}}^2}{2n}.$$

This Lemma says that if $\boldsymbol{S}$ is sufficiently close to $\boldsymbol{Z}$, then $\boldsymbol{Z}^\top \boldsymbol{S}$ is approximately an identity.

Proof Note that

$$\|\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q}\|_F^2 = 2nd - 2\left\langle \boldsymbol{Q}, \boldsymbol{Z}^\top \boldsymbol{S}\right\rangle$$

where $\|\boldsymbol{S}\|_F^2 = \|\boldsymbol{Z}\boldsymbol{Q}\|_F^2 = nd$. Note that

$$|\langle \boldsymbol{Q}, \boldsymbol{Z}^\top \boldsymbol{S}\rangle| \le \|\boldsymbol{Q}\|_{\mathrm{op}} \cdot \|\boldsymbol{Z}^\top \boldsymbol{S}\|_* = \|\boldsymbol{Z}^\top \boldsymbol{S}\|_*$$

where $\|\boldsymbol{Z}^\top \boldsymbol{S}\|_*$ denotes the nuclear norm of $\boldsymbol{Z}^\top \boldsymbol{S} \in \mathbb{R}^{d \times p}$. The maximum is assumed if $\boldsymbol{Q} = \boldsymbol{U}\boldsymbol{V}^\top$ where $\boldsymbol{U} \in \mathbb{R}^{d \times d}$ and $\boldsymbol{V} \in \mathbb{R}^{p \times d}$ are the left and right singular vectors of $\boldsymbol{Z}^\top \boldsymbol{S}$.

As a result, we get

$$d_F^2(\boldsymbol{S}, \boldsymbol{Z}) = 2\left(nd - \left\|\boldsymbol{Z}^\top \boldsymbol{S}\right\|_*\right) \le \frac{\delta^2 d \|\boldsymbol{\Delta}\|_{\mathrm{op}}^2}{n}.$$

Note that the largest singular value of $\boldsymbol{Z}^\top \boldsymbol{S}$ is at most n which trivially follows from triangle inequality. For the smallest singular value of $\boldsymbol{Z}^\top \boldsymbol{S}$, we use the following inequality

$$n - \sigma_{\min}\left(\boldsymbol{Z}^\top \boldsymbol{S}\right) \le \sum_{i=1}^{d}\left(n - \sigma_i\left(\boldsymbol{Z}^\top \boldsymbol{S}\right)\right) = nd - \left\|\boldsymbol{Z}^\top \boldsymbol{S}\right\|_* \le \frac{\delta^2 d \|\boldsymbol{\Delta}\|_{\mathrm{op}}^2}{2n}$$

which implies $\sigma_{\min}(\boldsymbol{Z}^\top \boldsymbol{S}) \ge n - 2\delta^2 n^{-1} d \|\boldsymbol{\Delta}\|_{\mathrm{op}}^2$. □

Lemma 5 *Suppose a second-order critical point* $\boldsymbol{S}$ *satisfies the proximity condition. Then*

$$\lambda_{\min}(\boldsymbol{\Lambda}_{ii}) \ge n - \frac{\delta^2 d \|\boldsymbol{\Delta}\|_{\mathrm{op}}^2}{2n} - \max_{1\le i\le n}\left\|\sum_{j\ne i}\boldsymbol{\Delta}_{ij}\boldsymbol{S}_j\right\|_{\mathrm{op}},$$

$$\max_{1\le i\le n}\left\|\sum_{j\ne i}\boldsymbol{\Delta}_{ij}\boldsymbol{S}_j\right\|_{\mathrm{op}} \le \delta\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}} \max_{1\le i\le n}\|\boldsymbol{\Delta}_i\|_{\mathrm{op}} + \max_{1\le i\le n}\left\|\boldsymbol{\Delta}_i^\top \boldsymbol{Z}\right\|_{\mathrm{op}}$$

where $\boldsymbol{\Delta}_i$ *is the* i*th block column of* $\boldsymbol{\Delta}$*.*

Proof Suppose $\boldsymbol{S}$ is a SOCP with $d_F(\boldsymbol{S}, \boldsymbol{Z}) \le \delta\sqrt{n^{-1}d}\|\boldsymbol{\Delta}\|_{\mathrm{op}}$. We have

$$\sigma_{\min}\left(\boldsymbol{Z}^\top \boldsymbol{S}\right) \ge n - \frac{\delta^2 d \|\boldsymbol{\Delta}\|_{\mathrm{op}}^2}{2n}, \quad \boldsymbol{\Lambda}_{ii} = \frac{1}{2}\sum_{j=1}^{n}\left(\boldsymbol{S}_i\boldsymbol{S}_j^\top \boldsymbol{A}_{ji} + \boldsymbol{A}_{ij}\boldsymbol{S}_j\boldsymbol{S}_i^\top\right) \succeq 0$$

from Lemma 4 and 3. The first order necessary condition (19) implies

$$\sum_{j=1}^{n}\boldsymbol{A}_{ij}\boldsymbol{S}_j = \boldsymbol{\Lambda}_{ii}\boldsymbol{S}_i, \qquad \boldsymbol{\Lambda}_{ii} = \sum_{j=1}^{n}\boldsymbol{S}_i\boldsymbol{S}_j^\top \boldsymbol{A}_{ji} = \sum_{j=1}^{n}\boldsymbol{A}_{ij}\boldsymbol{S}_j\boldsymbol{S}_i^\top$$

where $\boldsymbol{S}_i\boldsymbol{S}_i^\top = \boldsymbol{I}_d$. Therefore, the singular values of $\sum_{j=1}^{n}\boldsymbol{A}_{ij}\boldsymbol{S}_j$ and $\boldsymbol{\Lambda}_{ii}$ are the same. Moreover, due to the symmetry and $\boldsymbol{\Lambda}_{ii} \succeq 0$, its eigenvalues and singular values match:

$$\begin{aligned}
\lambda_{\min}(\boldsymbol{\Lambda}_{ii}) = \sigma_{\min}(\boldsymbol{\Lambda}_{ii}) &= \sigma_{\min}\left(\sum_{j=1}^{n}\boldsymbol{A}_{ij}\boldsymbol{S}_j\right) \\
&= \sigma_{\min}\left(\sum_{j=1}^{n}\boldsymbol{S}_j + \sum_{j=1}^{n}\boldsymbol{\Delta}_{ij}\boldsymbol{S}_j\right) \\
&\ge \sigma_{\min}\left(\boldsymbol{Z}^\top \boldsymbol{S}\right) - \left\|\sum_{j\ne i}\boldsymbol{\Delta}_{ij}\boldsymbol{S}_j\right\|_{\mathrm{op}}
\end{aligned}$$

$$\geq n - \frac{\delta^2 d \|\boldsymbol{\Delta}\|_{\mathrm{op}}^2}{2n} - \max_{1\leq i\leq n} \left\| \sum_{j\neq i} \boldsymbol{\Delta}_{ij} \boldsymbol{S}_j \right\|_{\mathrm{op}}$$

where the lower bound is independent of i. The key is to bound $\left\|\sum_{j\neq i} \boldsymbol{\Delta}_{ij}\boldsymbol{S}_j\right\|_{\mathrm{op}}$ which is suboptimal in this analysis.

$$\begin{aligned} \left\| \sum_{j\neq i} \boldsymbol{\Delta}_{ij} \boldsymbol{S}_j \right\|_{\mathrm{op}} &\leq \left\| \sum_{j\neq i} \boldsymbol{\Delta}_{ij} (\boldsymbol{S}_j - \boldsymbol{Q}) \right\|_{\mathrm{op}} + \left\| \sum_{j\neq i} \boldsymbol{\Delta}_{ij} \boldsymbol{Q} \right\|_{\mathrm{op}} \\ &\leq \|\boldsymbol{\Delta}_i^\top (\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q})\|_{\mathrm{op}} + \|\boldsymbol{\Delta}_i^\top \boldsymbol{Z}\|_{\mathrm{op}} \end{aligned}$$

where $\boldsymbol{\Delta}_i^\top \in \mathbb{R}^{d\times nd}$ is the ith row block of $\boldsymbol{\Delta}$. The operator norm of $\|\boldsymbol{\Delta}_i(\boldsymbol{S}-\boldsymbol{Z}\boldsymbol{Q})\|_{\mathrm{op}}$ is bounded by

$$\left\| \boldsymbol{\Delta}_i^\top (\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q}) \right\|_{\mathrm{op}} \leq \|\boldsymbol{\Delta}_i\|_{\mathrm{op}} \|\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q}\|_{\mathrm{op}} \leq \|\boldsymbol{\Delta}\|_{\mathrm{op}} \|\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q}\|_F.$$

Thus we have

$$\left\| \sum_{j\neq i} \boldsymbol{\Delta}_{ij} \boldsymbol{S}_j \right\|_{\mathrm{op}} \leq \|\boldsymbol{\Delta}_i\|_{\mathrm{op}} \|\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q}\|_F + \left\| \boldsymbol{\Delta}_i^\top \boldsymbol{Z} \right\|_{\mathrm{op}} \leq \delta \sqrt{\frac{d}{n}} \|\boldsymbol{\Delta}\|_{\mathrm{op}} \|\boldsymbol{\Delta}_i\|_{\mathrm{op}} + \left\| \boldsymbol{\Delta}_i^\top \boldsymbol{Z} \right\|_{\mathrm{op}}.$$

Taking the maximum over $1 \leq i \leq n$ gives the desired result. □

With this supporting lemma, we are ready to prove Proposition 2.

Proof of Proposition 2 The proof consists of two steps: first to show that $\boldsymbol{S}$ is a global maximizer to (BM) by showing that $\boldsymbol{\Lambda} - \boldsymbol{A} \succeq 0$; then prove that $\boldsymbol{S}$ is exactly rank-d.

Step 1 show that $\boldsymbol{S}$ is a global maximizer

Remember that $\boldsymbol{C}\boldsymbol{S} = 0$ if $\boldsymbol{S} \in \mathbb{R}^{nd\times p}$ is a critical point of f. Thus, to show $\boldsymbol{C}$ is positive semidefinite at critical point $\boldsymbol{S}$, it suffices to test $\boldsymbol{u}^\top \boldsymbol{C}\boldsymbol{u} \geq 0$ for all $\boldsymbol{u} \in \mathbb{R}^{nd\times 1}$ which is perpendicular to each column of $\boldsymbol{S}$:

$$\boldsymbol{S}^\top \boldsymbol{u} = 0 \in \mathbb{R}^{p\times 1} \iff \sum_{j=1}^n \boldsymbol{S}_j^\top \boldsymbol{u}_j = 0 \in \mathbb{R}^p$$

where $\boldsymbol{u}_j \in \mathbb{R}^d$ is the jth block of $\boldsymbol{u}$, $1 \leq j \leq n$.

Then it holds that

$$\begin{aligned} \boldsymbol{u}^\top \boldsymbol{C}\boldsymbol{u} &= \boldsymbol{u}^\top \boldsymbol{\Lambda}\boldsymbol{u} - \boldsymbol{u}^\top \boldsymbol{A}\boldsymbol{u} \\ &\geq \lambda_{\min}(\boldsymbol{\Lambda})\|\boldsymbol{u}\|^2 - \boldsymbol{u}^\top \left(\boldsymbol{Z}\boldsymbol{Z}^\top + \boldsymbol{\Delta}\right)\boldsymbol{u} \\ &\geq \lambda_{\min}(\boldsymbol{\Lambda})\|\boldsymbol{u}\|^2 - \boldsymbol{u}^\top \boldsymbol{Z}\boldsymbol{Z}^\top \boldsymbol{u} - \|\boldsymbol{\Delta}\|_{\mathrm{op}}\|\boldsymbol{u}\|^2 \end{aligned}$$

Note that $\lambda_{\min}(\boldsymbol{\Lambda}) = \min_{1\le i\le n}\lambda_{\min}(\boldsymbol{\Lambda}_{ii})$ which is given by Lemma 5. For $\boldsymbol{u}^\top \boldsymbol{Z}\boldsymbol{Z}^\top \boldsymbol{u}$, we use $\sum_{j=1}^n \boldsymbol{S}_j^\top \boldsymbol{u}_j = 0$ and

$$\begin{aligned}
\boldsymbol{u}^\top \boldsymbol{Z}\boldsymbol{Z}^\top \boldsymbol{u} &= \left\|\sum_{j=1}^n \boldsymbol{u}_j\right\|^2 = \left\|\sum_{j=1}^n \boldsymbol{Q}^\top \boldsymbol{u}_j\right\|^2 \\
&= \left\|\sum_{j=1}^n (\boldsymbol{S}_j - \boldsymbol{Q})^\top \boldsymbol{u}_j\right\|^2 \\
&= \left\|(\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q})^\top \boldsymbol{u}\right\|^2 \\
&\le \|\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q}\|_{\mathrm{op}}^2 \|\boldsymbol{u}\|^2 \\
&\le \frac{\delta^2 d\|\boldsymbol{\Delta}\|_{\mathrm{op}}^2}{n}\|\boldsymbol{u}\|^2
\end{aligned}$$

where the last inequality uses the proximity condition. For $\lambda_{\min}(\boldsymbol{\Lambda})$, we apply Lemma 5 and immediately arrive at:

$$\lambda_{\min}(\boldsymbol{C}) \ge n - \left(\frac{3\delta^2 d\|\boldsymbol{\Delta}\|_{\mathrm{op}}^2}{2n} + \delta\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}} \max_{1\le i\le n}\|\boldsymbol{\Delta}_i\|_{\mathrm{op}} + \max_{1\le i\le n}\left\|\boldsymbol{\Delta}_i^\top \boldsymbol{Z}\right\|_{\mathrm{op}} + \|\boldsymbol{\Delta}\|_{\mathrm{op}}\right) \ge 0.$$

Step 2 $\boldsymbol{S}$ is exactly rank-d We have shown the solution to the Burer-Monteiro approach is equivalent to that of the SDP. Now, we will prove that the solution to the Burer-Monteiro approach is exactly rank d.

How to show that $\boldsymbol{C}$ is rank-d deficient? It suffices to bound the dimension of its null space. In a more compact version, we have

$$\boldsymbol{C} = \boldsymbol{\Lambda} - \boldsymbol{A} = \boldsymbol{\Lambda} - \boldsymbol{Z}\boldsymbol{Z}^\top - \boldsymbol{\Delta}$$

The null space is bounded by

$$\begin{aligned}
\operatorname{null}(\boldsymbol{C}) &= nd - \operatorname{rank}\left(\boldsymbol{\Lambda} - \boldsymbol{Z}\boldsymbol{Z}^\top - \boldsymbol{\Delta}\right) \\
&\le nd + d - \operatorname{rank}(\boldsymbol{\Lambda} - \boldsymbol{\Delta})
\end{aligned}$$

where

$$\operatorname{rank}(\boldsymbol{\Lambda} - \boldsymbol{Z}\boldsymbol{Z}^\top - \boldsymbol{\Delta}) + \operatorname{rank}(\boldsymbol{Z}\boldsymbol{Z}^\top) = \operatorname{rank}(\boldsymbol{\Lambda} - \boldsymbol{Z}\boldsymbol{Z}^\top - \boldsymbol{\Delta}) + d \ge \operatorname{rank}(\boldsymbol{\Lambda} - \boldsymbol{\Delta}).$$

It suffices to provide a lower bound of $\operatorname{rank}(\boldsymbol{\Lambda} - \boldsymbol{\Delta})$. In particular, we aim to show that $\boldsymbol{\Lambda} - \boldsymbol{\Delta}$ is full-rank by

$$\boldsymbol{\Lambda} - \boldsymbol{\Delta} \succ 0.$$

This is guaranteed by

$$\lambda_{\min}(\boldsymbol{\Lambda}) > \|\boldsymbol{\Delta}\|_{\text{op}}$$

and more explicitly

$$\begin{aligned}
&\lambda_{\min}(\boldsymbol{\Lambda}) - \|\boldsymbol{\Delta}\|_{\text{op}} \\
&\geq n - \left(\frac{\delta^2 d \|\boldsymbol{\Delta}\|_{\text{op}}^2}{2n} + \delta \sqrt{\frac{d}{n}} \|\boldsymbol{\Delta}\|_{\text{op}} \max_{1\leq i\leq n} \|\boldsymbol{\Delta}_i\|_{\text{op}} + \max_{1\leq i\leq n} \left\| \boldsymbol{\Delta}_i^\top \boldsymbol{Z} \right\|_{\text{op}} + \|\boldsymbol{\Delta}\|_{\text{op}} \right) > 0.
\end{aligned}$$

Then we have $\text{null}(\boldsymbol{C}) \leq nd + d - \text{rank}(\boldsymbol{\Lambda} - \boldsymbol{\Delta}) = nd + d - nd = d$. If $\boldsymbol{C}$ is of rank $nd - d$, then $\text{rank}(\boldsymbol{S}) = d$. Thus the global optimum is the same as that of (SDP) and (P). □

5.5 Proof of Proposition 3

Proof of Proposition 3 It is unclear how to characterize the global maximizer to the objective function (BM). However, the global maximizer must be a 2nd critical point whose corresponding objective function value is greater than $f(\boldsymbol{S})$ evaluated at the fully synchronous state $\boldsymbol{S}_i = \boldsymbol{S}_j$.

Throughout our discussion, we let $\boldsymbol{Q}$ be the minimizer to $\min_{\boldsymbol{Q}\in\text{St}(d,p)} \|\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q}\|_F$. Given $\boldsymbol{S}$ which satisfies $f(\boldsymbol{S}) \geq f(\boldsymbol{Z}\boldsymbol{Q})$, we have

$$f(\boldsymbol{S}) \geq f(\boldsymbol{Z}\boldsymbol{Q}) \iff \left\langle \boldsymbol{Z}\boldsymbol{Z}^\top + \boldsymbol{\Delta}, \boldsymbol{S}\boldsymbol{S}^\top \right\rangle \geq \left\langle \boldsymbol{Z}\boldsymbol{Z}^\top + \boldsymbol{\Delta}, \boldsymbol{Z}\boldsymbol{Z}^\top \right\rangle.$$

Note that $\langle \boldsymbol{Z}\boldsymbol{Z}^\top, \boldsymbol{Z}\boldsymbol{Z}^\top \rangle = n^2 d$ and $\langle \boldsymbol{Z}\boldsymbol{Z}^\top, \boldsymbol{S}\boldsymbol{S}^\top \rangle = \|\boldsymbol{Z}^\top \boldsymbol{S}\|_F^2$. This gives

$$\begin{aligned}
n^2 d - \|\boldsymbol{Z}^\top \boldsymbol{S}\|_F^2 &\leq \left\langle \boldsymbol{\Delta}, \boldsymbol{S}\boldsymbol{S}^\top - \boldsymbol{Z}\boldsymbol{Z}^\top \right\rangle \\
\text{(Due to the symmetry of } \boldsymbol{\Delta}) \quad &= \left\langle \boldsymbol{\Delta}, \boldsymbol{S}\boldsymbol{S}^\top - \boldsymbol{Z}\boldsymbol{Q}\boldsymbol{S}^\top + \boldsymbol{S}\boldsymbol{Q}^\top \boldsymbol{Z}^\top + \boldsymbol{Z}\boldsymbol{Z}^\top \right\rangle \\
&= \left\langle \boldsymbol{\Delta}, (\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q})(\boldsymbol{S} + \boldsymbol{Z}\boldsymbol{Q})^\top \right\rangle \\
&\leq \|\boldsymbol{\Delta}(\boldsymbol{S} + \boldsymbol{Z}\boldsymbol{Q})\|_F \cdot d_F(\boldsymbol{S}, \boldsymbol{Z}) \qquad (24)
\end{aligned}$$

where $d_F(\boldsymbol{S}, \boldsymbol{Z}) = \|\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q}\|_F$ and

$$(\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q})(\boldsymbol{S} + \boldsymbol{Z}\boldsymbol{Q})^\top = \boldsymbol{S}\boldsymbol{S}^\top - (\boldsymbol{Z}\boldsymbol{Q})^\top \boldsymbol{S} + \boldsymbol{S}(\boldsymbol{Z}\boldsymbol{Q})^\top - \boldsymbol{Z}\boldsymbol{Z}^\top$$

In fact, $n^2 d - \|\boldsymbol{Z}^\top \boldsymbol{S}\|_F^2$ is well controlled by $d_F(\boldsymbol{S}, \boldsymbol{Z})$:

$$n^2 d - \left\| \boldsymbol{Z}^\top \boldsymbol{S} \right\|_F^2 = \sum_{i=1}^d \left(n^2 - \sigma_i^2 \left(\boldsymbol{Z}^\top \boldsymbol{S} \right) \right)$$

$$= \sum_{i=1}^{d} \left(n + \sigma_i\left(\boldsymbol{Z}^\top \boldsymbol{S}\right)\right)\left(n - \sigma_i\left(\boldsymbol{Z}^\top \boldsymbol{S}\right)\right) \tag{25}$$

where $\sigma_i(\boldsymbol{Z}^\top \boldsymbol{S})$ is the ith largest singular value of $\boldsymbol{Z}^\top \boldsymbol{S}$. On the other hand, it holds that

$$\frac{1}{2}\min_{\boldsymbol{Q}\in \mathrm{St}(d,p)} \|\boldsymbol{S} - \boldsymbol{Z}\boldsymbol{Q}\|_F^2 = nd - \left\|\boldsymbol{Z}^\top \boldsymbol{S}\right\|_* = \sum_{i=1}^{d}\left(n - \sigma_i\left(\boldsymbol{Z}^\top \boldsymbol{S}\right)\right)$$

Remember that $0 \le \sigma_i(\boldsymbol{Z}^\top \boldsymbol{S}) \le n$ due to the orthogonality of each $\boldsymbol{S}_i$. Therefore, we have

$$\begin{aligned}\frac{nd_F^2(\boldsymbol{S},\boldsymbol{Z})}{2} &\le n\left(nd - \|\boldsymbol{Z}^\top \boldsymbol{S}\|_*\right) \le n^2 d - \left\|\boldsymbol{Z}^\top \boldsymbol{S}\right\|_F^2 \\ &\le 2n\left(nd - \|\boldsymbol{Z}^\top \boldsymbol{S}\|_*\right) \le nd_F^2(\boldsymbol{S},\boldsymbol{Z}) \end{aligned}\tag{26}$$

which follows from (25). Substitute it back into (24), and we get

$$\frac{nd^2(\boldsymbol{S},\boldsymbol{Z})}{2} \le n^2 d - \left\|\boldsymbol{Z}^\top \boldsymbol{S}\right\|_F^2 \le \|\boldsymbol{\Delta}(\boldsymbol{S} + \boldsymbol{Z}\boldsymbol{Q})\|_F \cdot d(\boldsymbol{S},\boldsymbol{Z})$$

Immediately, we have the following estimate of $d(\boldsymbol{S},\boldsymbol{Z})$:

$$\begin{aligned} d_F(\boldsymbol{S},\boldsymbol{Z}) &\le \frac{2}{n}\|\boldsymbol{\Delta}(\boldsymbol{S} + \boldsymbol{Z}\boldsymbol{Q})\|_F \\ &\le \frac{2}{n}\cdot \|\boldsymbol{\Delta}\|_{\mathrm{op}}\|\boldsymbol{S} + \boldsymbol{Z}\boldsymbol{Q}\|_F \\ &\le \frac{2}{n}\cdot \|\boldsymbol{\Delta}\|_{\mathrm{op}} \cdot 2\sqrt{nd} \\ &\le 4\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}} \end{aligned}$$

where $\|\boldsymbol{S} + \boldsymbol{Z}\boldsymbol{Q}\|_F \le 2\sqrt{nd}$ follows from $\|\boldsymbol{S}\|_F = \|\boldsymbol{Z}\|_F = \sqrt{nd}$. □

5.6 Proof of Proposition 4

This section is devoted to proving all the SOCPs are highly aligned with the fully synchronized state. The proof follows from two steps: (a) using the second order necessary condition to show that all SOCPs have a large objective function value; (b) combining (a) with the first order necessary condition leads to Proposition 4.

Lemma 6 *All the second order critical points $\boldsymbol{S} \in \mathrm{St}(d,p)^{\otimes n}$ must satisfy:*

$$(p-d)\left\|\boldsymbol{Z}^\top \boldsymbol{S}\right\|_F^2 \ge (p-2d)n^2 d + \left\|\boldsymbol{S}\boldsymbol{S}^\top\right\|_F^2 d$$

$$+\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\left\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\right\|_F^2-d\right)\operatorname{Tr}(\boldsymbol{\Delta}_{ij})+(p-d)\left\langle\boldsymbol{\Delta},\boldsymbol{Z}\boldsymbol{Z}^\top-\boldsymbol{S}\boldsymbol{S}^\top\right\rangle.$$

Suppose the noise is zero, then $(p-d)\|\boldsymbol{Z}^\top\boldsymbol{S}\|_F^2\geq(p-2d)n^2d+\|\boldsymbol{S}\boldsymbol{S}^\top\|_F^2 d$ holds. It means that $\|\boldsymbol{Z}^\top\boldsymbol{S}\|_F^2$ is quite close to n^2d, i.e., $\{\boldsymbol{S}_i\}_{i=1}^n$ are highly aligned, if p is reasonably large. The proof idea of this lemma can also be found in [40, 41].

Proof Let's first consider the second-order necessary condition (20):

$$\sum_{i=1}^{n}\left\langle\boldsymbol{\Lambda}_{ii},\dot{\boldsymbol{S}}_i\dot{\boldsymbol{S}}_i^\top\right\rangle\geq\sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle\boldsymbol{A}_{ij},\dot{\boldsymbol{S}}_i\dot{\boldsymbol{S}}_j^\top\right\rangle$$

for all $\dot{\boldsymbol{S}}_i$ on the tangent space of $\mathrm{St}(p,d)$ at $\boldsymbol{S}_i$ where $\boldsymbol{\Lambda}_{ii}=\frac{1}{2}\sum_{j=1}^{n}(\boldsymbol{S}_i\boldsymbol{S}_j^\top\boldsymbol{A}_{ji}+\boldsymbol{A}_{ij}\boldsymbol{S}_j\boldsymbol{S}_i^\top)$. Now we pick $\dot{\boldsymbol{S}}_i$ as

$$\dot{\boldsymbol{S}}_i=\boldsymbol{\Phi}\left(\boldsymbol{I}_p-\boldsymbol{S}_i^\top\boldsymbol{S}_i\right)\in\mathbb{R}^{d\times p}$$

where $\boldsymbol{\Phi}\in\mathbb{R}^{d\times p}$ is a Gaussian random matrix, i.e., each entry in $\boldsymbol{\Phi}$ is an i.i.d. $\mathcal{N}(0,1)$ random variable. It is easy to verify that $\dot{\boldsymbol{S}}_i$ is indeed on the tangent space since $\boldsymbol{S}_i\dot{\boldsymbol{S}}_i^\top=0$. By taking the expectation w.r.t. $\boldsymbol{\Phi}$, the inequality still holds:

$$\sum_{i=1}^{n}\left\langle\boldsymbol{\Lambda}_{ii},\mathbb{E}\,\dot{\boldsymbol{S}}_i\dot{\boldsymbol{S}}_i^\top\right\rangle\geq\sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle\boldsymbol{A}_{ij},\mathbb{E}\,\dot{\boldsymbol{S}}_i\dot{\boldsymbol{S}}_j^\top\right\rangle.$$

It suffices to compute $\mathbb{E}\,\dot{\boldsymbol{S}}_i\dot{\boldsymbol{S}}_j^\top$ now.

$$\begin{aligned}\mathbb{E}\left(\dot{\boldsymbol{S}}_i\dot{\boldsymbol{S}}_j^\top\right)&=\mathbb{E}\,\boldsymbol{\Phi}\left(\boldsymbol{I}_p-\boldsymbol{S}_i^\top\boldsymbol{S}_i\right)\left(\boldsymbol{I}_p-\boldsymbol{S}_j^\top\boldsymbol{S}_j\right)\boldsymbol{\Phi}^\top\\&=\left\langle\boldsymbol{I}_p-\boldsymbol{S}_i^\top\boldsymbol{S}_i,\boldsymbol{I}_p-\boldsymbol{S}_j^\top\boldsymbol{S}_j\right\rangle\boldsymbol{I}_d\\&=\left(p-2d+\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\|_F^2\right)\boldsymbol{I}_d\\&=\begin{cases}(p-d)\boldsymbol{I}_d, & i=j,\\ \left(p-2d+\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\|_F^2\right)\boldsymbol{I}_d, & i\neq j\end{cases}\end{aligned}$$

where $d=\operatorname{Tr}(\boldsymbol{S}_i\boldsymbol{S}_i^\top)$. Therefore, we have

$$(p-d)\sum_{i=1}^{n}\operatorname{Tr}(\boldsymbol{\Lambda}_{ii})\geq\sum_{i=1}^{n}\sum_{j=1}^{n}\left(p-2d+\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\|_F^2\right)\operatorname{Tr}(\boldsymbol{A}_{ij}).\tag{27}$$

The right hand side of (27) equals

$$\begin{aligned}
&\sum_{i=1}^{n}\sum_{j=1}^{n}\left(p-2d+\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\|_F^2\right)\mathrm{Tr}(\boldsymbol{A}_{ij})\\
&\quad=\sum_{i=1}^{n}\sum_{j=1}^{n}\left(p-2d+\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\|_F^2\right)\left(d+\mathrm{Tr}\left(\boldsymbol{\Delta}_{ij}\right)\right)\\
&\quad=(p-2d)n^2d+\left\|\boldsymbol{S}\boldsymbol{S}^\top\right\|_F^2 d+(p-2d)\left\langle\boldsymbol{\Delta},\boldsymbol{Z}\boldsymbol{Z}^\top\right\rangle+\sum_{i=1}^{n}\sum_{j=1}^{n}\left\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\right\|_F^2\mathrm{Tr}(\boldsymbol{\Delta}_{ij})
\end{aligned}$$

where $\boldsymbol{A}_{ij}=\boldsymbol{I}_d+\boldsymbol{\Delta}_{ij}$. From the definition of $\boldsymbol{\Lambda}_{ii}$, the left side of (27) equal to

$$\begin{aligned}
\sum_{i=1}^{n}\mathrm{Tr}(\boldsymbol{\Lambda}_{ii})&=\sum_{i=1}^{n}\sum_{j=1}^{n}\left\langle\boldsymbol{A}_{ij},\boldsymbol{S}_i\boldsymbol{S}_j^\top\right\rangle\\
&=\left\langle\boldsymbol{A},\boldsymbol{S}\boldsymbol{S}^\top\right\rangle\\
&=\left\langle\boldsymbol{Z}\boldsymbol{Z}^\top+\boldsymbol{\Delta},\boldsymbol{S}\boldsymbol{S}^\top\right\rangle\\
&=\left\|\boldsymbol{Z}^\top\boldsymbol{S}\right\|_F^2+\left\langle\boldsymbol{\Delta},\boldsymbol{S}\boldsymbol{S}^\top\right\rangle.
\end{aligned}$$

Plugging the estimation back to (27) results in

$$\begin{aligned}
(p-d)\left(\left\|\boldsymbol{Z}^\top\boldsymbol{S}\right\|_F^2+\left\langle\boldsymbol{\Delta},\boldsymbol{S}\boldsymbol{S}^\top\right\rangle\right)&\geq(p-2d)n^2d+\left\|\boldsymbol{S}\boldsymbol{S}^\top\right\|_F^2 d\\
&\quad+(p-2d)\left\langle\boldsymbol{\Delta},\boldsymbol{Z}\boldsymbol{Z}^\top\right\rangle+\sum_{i=1}^{n}\sum_{j=1}^{n}\left\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\right\|_F^2\mathrm{Tr}(\boldsymbol{\Delta}_{ij}).
\end{aligned}$$

By separating the signal from the noise, we have

$$\begin{aligned}
(p-d)\left\|\boldsymbol{Z}^\top\boldsymbol{S}\right\|_F^2&\geq(p-2d)n^2d+\left\|\boldsymbol{S}\boldsymbol{S}^\top\right\|_F^2 d\\
&\quad+\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\left\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\right\|_F^2-d\right)\mathrm{Tr}\left(\boldsymbol{\Delta}_{ij}\right)+(p-d)\left\langle\boldsymbol{\Delta},\boldsymbol{Z}\boldsymbol{Z}^\top-\boldsymbol{S}\boldsymbol{S}^\top\right\rangle
\end{aligned}$$

where $\langle\boldsymbol{\Delta},\boldsymbol{Z}\boldsymbol{Z}^\top\rangle=\sum_{i=1}^n\sum_{j=1}^n\mathrm{Tr}(\boldsymbol{\Delta}_{ij})$. □

Lemma 7 *Any first-order critical point satisfies:*

$$\|\boldsymbol{S}\boldsymbol{S}^\top\|_F^2\geq\left\|\boldsymbol{Z}^\top\boldsymbol{S}\right\|_F^2-\frac{1}{n}\|\boldsymbol{\Delta}\boldsymbol{S}\|_F^2.$$

Proof Note that the first-order necessary condition is

$$\sum_{j=1}^{n} \boldsymbol{A}_{ij}\boldsymbol{S}_j - \frac{1}{2}\sum_{j=1}^{n}\left(\boldsymbol{S}_i\boldsymbol{S}_j^\top \boldsymbol{A}_{ji} + \boldsymbol{A}_{ij}\boldsymbol{S}_j\boldsymbol{S}_i^\top\right)\boldsymbol{S}_i = 0$$

which implies $\sum_{j=1} \boldsymbol{S}_i\boldsymbol{S}_j^\top \boldsymbol{A}_{ji} = \sum_{j=1} \boldsymbol{A}_{ij}\boldsymbol{S}_j\boldsymbol{S}_i^\top$ by applying $\boldsymbol{S}_i^\top$ to the equation above. Then it reduces to

$$\sum_{j=1}^{n} \boldsymbol{A}_{ij}\boldsymbol{S}_j\left(\boldsymbol{I}_p - \boldsymbol{S}_i^\top\boldsymbol{S}_i\right) = 0 \iff \left(\boldsymbol{Z}^\top\boldsymbol{S} + \sum_{j=1}^{n}\boldsymbol{\Delta}_{ij}\boldsymbol{S}_j\right)\left(\boldsymbol{I}_p - \boldsymbol{S}_i^\top\boldsymbol{S}_i\right) = 0.$$

By separating the signal from the noise, we have

$$\boldsymbol{Z}^\top\boldsymbol{S}\left(\boldsymbol{I}_p - \boldsymbol{S}_i^\top\boldsymbol{S}_i\right) = -\sum_{j=1}^{n}\boldsymbol{\Delta}_{ij}\boldsymbol{S}_j\left(\boldsymbol{I}_p - \boldsymbol{S}_i^\top\boldsymbol{S}_i\right)$$

Taking the Frobenius norm leads to

$$\left\|\boldsymbol{Z}^\top\boldsymbol{S}\right\|_F^2 - \left\langle \boldsymbol{Z}^\top\boldsymbol{S}\boldsymbol{S}_i^\top\boldsymbol{S}_i, \boldsymbol{Z}^\top\boldsymbol{S}\right\rangle \le \left\|\sum_{j=1}^{n}\boldsymbol{\Delta}_{ij}\boldsymbol{S}_j\right\|_F^2$$

where $\|\boldsymbol{I}_p - \boldsymbol{S}_i^\top\boldsymbol{S}_i\|_{\mathrm{op}} = 1$. Taking the sum over $1 \le i \le n$ gives

$$\begin{aligned}
\|\boldsymbol{\Delta}\boldsymbol{S}\|_F^2 &= \sum_{i=1}^{n}\left\|\sum_{j=1}^{n}\boldsymbol{\Delta}_{ij}\boldsymbol{S}_j\right\|_F^2 \\
&\ge n\left\|\boldsymbol{Z}^\top\boldsymbol{S}\right\|_F^2 - \sum_{i=1}^{n}\left\langle \boldsymbol{Z}^\top\boldsymbol{S}\boldsymbol{S}_i^\top\boldsymbol{S}_i, \boldsymbol{Z}^\top\boldsymbol{S}\right\rangle \\
&= n\left\|\boldsymbol{Z}^\top\boldsymbol{S}\right\|_F^2 - \left\langle \boldsymbol{Z}^\top\boldsymbol{S}\boldsymbol{S}^\top\boldsymbol{S}, \boldsymbol{Z}^\top\boldsymbol{S}\right\rangle \\
&= n\left\|\boldsymbol{Z}^\top\boldsymbol{S}\right\|_F^2 - \left\langle \boldsymbol{S}\boldsymbol{S}^\top\boldsymbol{S}\boldsymbol{S}^\top, \boldsymbol{Z}\boldsymbol{Z}^\top, \right\rangle
\end{aligned}$$

where $\boldsymbol{S}^\top\boldsymbol{S} = \sum_{i=1}^{n}\boldsymbol{S}_i^\top\boldsymbol{S}_i$. Thus

$$\left\|\boldsymbol{Z}^\top\boldsymbol{S}\right\|_F^2 - \frac{1}{n}\|\boldsymbol{\Delta}\boldsymbol{S}\|_F^2 \le \frac{1}{n}\left\langle \boldsymbol{S}\boldsymbol{S}^\top\boldsymbol{S}\boldsymbol{S}^\top, \boldsymbol{Z}\boldsymbol{Z}^\top\right\rangle \le \left\|\boldsymbol{S}\boldsymbol{S}^\top\right\|_F^2$$

where $\boldsymbol{Z}\boldsymbol{Z}^\top = \boldsymbol{J}_n \otimes \boldsymbol{I}_d \preceq n\boldsymbol{I}_{nd}$. □

Lemma 8 *For any matrix $\boldsymbol{X} \in \mathbb{R}^{nd\times nd}$, it holds that*

$$\left\| \boldsymbol{X} \circ \boldsymbol{S}\boldsymbol{S}^\top \right\|_{\mathrm{op}} \leq \|\boldsymbol{X}\|_{\mathrm{op}}$$

where $\boldsymbol{S} \in \mathbb{R}^{p\times nd}$ and $[\boldsymbol{S}\boldsymbol{S}^\top]_{ii} = \boldsymbol{I}_d$. Here " $\circ$ " stands for the Hadamard product of two matrices.

Proof Let the $\boldsymbol{u}_\ell \in \mathbb{R}^{nd\times 1}$ be the ℓth column of $\boldsymbol{S}$ where $1 \leq \ell \leq p$ and let $\boldsymbol{\varphi} \in \mathbb{R}^{nd\times 1}$ be an arbitrary unit vector.

$$\begin{aligned}
\boldsymbol{\varphi}^\top \left(\boldsymbol{X} \circ \boldsymbol{S}\boldsymbol{S}^\top\right)\boldsymbol{\varphi} &= \sum_{\ell=1}^{p} \boldsymbol{\varphi}^\top \left(\boldsymbol{X} \circ \boldsymbol{u}_\ell \boldsymbol{u}_\ell^\top\right)\boldsymbol{\varphi} \\
&= \sum_{\ell=1}^{p} \boldsymbol{\varphi}^\top \operatorname{diag}(\boldsymbol{u}_\ell)\boldsymbol{X} \operatorname{diag}(\boldsymbol{u}_\ell)\boldsymbol{\varphi} \\
&\leq \|\boldsymbol{X}\|_{\mathrm{op}} \cdot \sum_{\ell=1}^{p} \|\operatorname{diag}(\boldsymbol{u}_\ell)\boldsymbol{\varphi}\|^2 \\
&= \|\boldsymbol{X}\|_{\mathrm{op}} \|\operatorname{diag}(\boldsymbol{\varphi})[\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p]\|_F^2 \\
&= \|\boldsymbol{X}\|_{\mathrm{op}} \cdot \|\operatorname{diag}(\boldsymbol{\varphi})\boldsymbol{S}\|_F^2 \\
&= \|\boldsymbol{X}\|_{\mathrm{op}} \cdot \operatorname{Tr}(\operatorname{diag}(\boldsymbol{\varphi})\boldsymbol{S}\boldsymbol{S}^\top \operatorname{diag}(\boldsymbol{\varphi})) \\
&= \|\boldsymbol{X}\|_{\mathrm{op}} \|\boldsymbol{\varphi}\|^2
\end{aligned}$$

where $\operatorname{diag}(\boldsymbol{S}\boldsymbol{S}^\top) = \boldsymbol{I}_{nd}$. Thus we have shown $\|\boldsymbol{X} \circ \boldsymbol{S}\boldsymbol{S}^\top\|_{\mathrm{op}} \leq \|\boldsymbol{X}\|_{\mathrm{op}}$. □

Proof of Proposition 4 Lemmas 6 and 7 imply that all the SOCPs of (BM) satisfy

$$\begin{aligned}
(p-d)\|\boldsymbol{Z}^\top\boldsymbol{S}\|_F^2 &\geq (p-2d)n^2 d + d\left\|\boldsymbol{Z}^\top\boldsymbol{S}\right\|_F^2 - \frac{d}{n}\|\boldsymbol{\Delta}\boldsymbol{S}\|_F^2 \\
&\quad + \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\left\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\right\|_F^2 - d\right)\operatorname{Tr}(\boldsymbol{\Delta}_{ij}) + (p-d)\left\langle \boldsymbol{\Delta}, \boldsymbol{Z}\boldsymbol{Z}^\top - \boldsymbol{S}\boldsymbol{S}^\top\right\rangle
\end{aligned}$$

where $\|\boldsymbol{S}\boldsymbol{S}^\top\|_F^2 \geq \left\|\boldsymbol{Z}^\top\boldsymbol{S}\right\|_F^2 - \frac{1}{n}\|\boldsymbol{\Delta}\boldsymbol{S}\|_F^2$. This is equivalent to

$$\begin{aligned}
(p-2d)\left(n^2 d - \|\boldsymbol{Z}^\top\boldsymbol{S}\|_F^2\right) &\leq \underbrace{\frac{d}{n}\|\boldsymbol{\Delta}\boldsymbol{S}\|_F^2}_{T_1} - \underbrace{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\|_F^2 - d\right)\operatorname{Tr}(\boldsymbol{\Delta}_{ij})}_{T_2} \\
&\quad + (p-d)\underbrace{\left\langle \boldsymbol{\Delta}, \boldsymbol{S}\boldsymbol{S}^\top - \boldsymbol{Z}\boldsymbol{Z}^\top\right\rangle}_{T_3} \\
&\leq |T_1| + |T_2| + (p-d)|T_3|.
\end{aligned}$$

Estimation of $|T_1|$ **and** $|T_3|$**:** For T_1, we simply have

$$|T_1| \le \frac{d}{n}\|\boldsymbol{\Delta}\|_{\text{op}}^2\|\boldsymbol{S}\|_F^2 = \frac{d}{n}\cdot\|\boldsymbol{\Delta}\|_{\text{op}}^2\cdot nd = d^2\|\boldsymbol{\Delta}\|_{\text{op}}^2.$$

For T_3, we have

$$\begin{aligned}|T_3| &= \left|\left\langle \boldsymbol{\Delta}, \boldsymbol{S}\boldsymbol{S}^\top - \boldsymbol{Z}\boldsymbol{Z}^\top\right\rangle\right| \\ &= \left|\left\langle \boldsymbol{\Delta}, (\boldsymbol{S}-\boldsymbol{Z}\boldsymbol{Q})\,(\boldsymbol{S}+\boldsymbol{Z}\boldsymbol{Q})^\top\right\rangle\right| \\ &\le \|\boldsymbol{\Delta}\|_{\text{op}}\cdot\|\boldsymbol{S}+\boldsymbol{Z}\boldsymbol{Q}\|_F\cdot\|\boldsymbol{S}-\boldsymbol{Z}\boldsymbol{Q}\|_F \\ &\le 2\|\boldsymbol{\Delta}\|_{\text{op}}\sqrt{nd}\cdot d_F(\boldsymbol{S},\boldsymbol{Z}).\end{aligned}$$

Estimation of $|T_2|$: Define a new matrix $\widetilde{\boldsymbol{\Delta}} := \operatorname{Tr}_d(\boldsymbol{\Delta})\otimes\boldsymbol{J}_d$ whose (i,j)-entry block is $\operatorname{Tr}(\boldsymbol{\Delta}_{ij})\boldsymbol{J}_d$ and $\|\widetilde{\boldsymbol{\Delta}}\|_{\text{op}} = \|\operatorname{Tr}_d(\boldsymbol{\Delta})\|_{\text{op}}d$. Note that

$$\begin{aligned}\sum_{i=1}^n\sum_{j=1}^n\left(\left\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\right\|_F^2 - d\right)\operatorname{Tr}(\boldsymbol{\Delta}_{ij}) &= \sum_{i=1}^n\sum_{j=1}^n\left\langle \boldsymbol{S}_i\boldsymbol{S}_j^\top\circ\boldsymbol{S}_i\boldsymbol{S}_j^\top - \boldsymbol{I}_d, \boldsymbol{J}_d\right\rangle\operatorname{Tr}(\boldsymbol{\Delta}_{ij}) \\ &= \sum_{i=1}^n\sum_{j=1}^n\left\langle \boldsymbol{S}_i\boldsymbol{S}_j^\top\circ\boldsymbol{S}_i\boldsymbol{S}_j^\top - \boldsymbol{I}_d, \widetilde{\boldsymbol{\Delta}}_{ij}\right\rangle \\ &= \left\langle \boldsymbol{S}\boldsymbol{S}^\top\circ\boldsymbol{S}\boldsymbol{S}^\top - \boldsymbol{Z}\boldsymbol{Z}^\top\circ\boldsymbol{Z}\boldsymbol{Z}^\top, \widetilde{\boldsymbol{\Delta}}\right\rangle \\ &= \left\langle\left(\boldsymbol{S}\boldsymbol{S}^\top - \boldsymbol{Z}\boldsymbol{Z}^\top\right)\circ\left(\boldsymbol{S}\boldsymbol{S}^\top + \boldsymbol{Z}\boldsymbol{Z}^\top\right), \widetilde{\boldsymbol{\Delta}}\right\rangle\end{aligned}$$

where $(\boldsymbol{S}\boldsymbol{S}^\top\circ\boldsymbol{S}\boldsymbol{S}^\top)_{ij} = \boldsymbol{S}_i\boldsymbol{S}_j^\top\circ\boldsymbol{S}_i\boldsymbol{S}_j^\top$, $(\boldsymbol{Z}\boldsymbol{Z}^\top\circ\boldsymbol{Z}\boldsymbol{Z}^\top)_{ij} = \boldsymbol{I}_d$, and $\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\|_F^2 = \langle\boldsymbol{S}_i\boldsymbol{S}_j^\top, \boldsymbol{S}_i\boldsymbol{S}_j^\top\rangle = \langle\boldsymbol{S}_i\boldsymbol{S}_j^\top\circ\boldsymbol{S}_i\boldsymbol{S}_j^\top, \boldsymbol{J}_d\rangle$. As a result, we have

$$\begin{aligned}\sum_{i=1}^n\sum_{j=1}^n\left(\left\|\boldsymbol{S}_i\boldsymbol{S}_j^\top\right\|_F^2 - d\right)\operatorname{Tr}(\boldsymbol{\Delta}_{ij}) &= \left\langle \boldsymbol{S}\boldsymbol{S}^\top - \boldsymbol{Z}\boldsymbol{Z}^\top, \widetilde{\boldsymbol{\Delta}}\circ\left(\boldsymbol{S}\boldsymbol{S}^\top + \boldsymbol{Z}\boldsymbol{Z}^\top\right)\right\rangle \\ &= \left\langle (\boldsymbol{S}-\boldsymbol{Z}\boldsymbol{Q})(\boldsymbol{S}+\boldsymbol{Z}\boldsymbol{Q})^\top, \widetilde{\boldsymbol{\Delta}}\circ\left(\boldsymbol{S}\boldsymbol{S}^\top + \boldsymbol{Z}\boldsymbol{Z}^\top\right)\right\rangle \\ &\le \|\boldsymbol{S}-\boldsymbol{Z}\boldsymbol{Q}\|_F\cdot\left\|\widetilde{\boldsymbol{\Delta}}\circ\left(\boldsymbol{S}\boldsymbol{S}^\top + \boldsymbol{Z}\boldsymbol{Z}^\top\right)\right\|_{\text{op}}\|\boldsymbol{S}+\boldsymbol{Z}\boldsymbol{Q}\|_F \\ &\le d_F(\boldsymbol{S},\boldsymbol{Z})\cdot 2\sqrt{nd}\left\|\widetilde{\boldsymbol{\Delta}}\circ\left(\boldsymbol{S}\boldsymbol{S}^\top + \boldsymbol{Z}\boldsymbol{Z}^\top\right)\right\|_{\text{op}}.\end{aligned}$$

Now the goal is to get an upper bound of $\|\widetilde{\boldsymbol{\Delta}}\circ(\boldsymbol{S}\boldsymbol{S}^\top + \boldsymbol{Z}\boldsymbol{Z}^\top)\|_{\text{op}}$. In fact, it holds

$$\begin{aligned}\|\widetilde{\boldsymbol{\Delta}}\circ(\boldsymbol{S}\boldsymbol{S}^\top + \boldsymbol{Z}\boldsymbol{Z}^\top)\|_{\text{op}} &\le \|\widetilde{\boldsymbol{\Delta}}\circ\boldsymbol{S}\boldsymbol{S}^\top\|_{\text{op}} + \|\widetilde{\boldsymbol{\Delta}}\circ\boldsymbol{Z}\boldsymbol{Z}^\top\|_{\text{op}} \\ &\le 2\|\widetilde{\boldsymbol{\Delta}}\|_{\text{op}} = 2d\|\operatorname{Tr}_d(\boldsymbol{\Delta})\|_{\text{op}}\end{aligned}$$

where the second inequality follows from Lemma 8 and $\|\widetilde{\boldsymbol{\Delta}}\|_{\mathrm{op}} = d\|\operatorname{Tr}_d(\boldsymbol{\Delta})\|_{\mathrm{op}}$. Therefore,

$$\begin{aligned}|T_2| &\le d_F(\boldsymbol{S},\boldsymbol{Z})\cdot 2\sqrt{nd}\left\|\widetilde{\boldsymbol{\Delta}}\circ\left(\boldsymbol{S}\boldsymbol{S}^\top+\boldsymbol{Z}\boldsymbol{Z}^\top\right)\right\|_{\mathrm{op}}\\ &\le d_F(\boldsymbol{S},\boldsymbol{Z})\cdot 2\sqrt{nd}\cdot 2d\,\|\operatorname{Tr}_d(\boldsymbol{\Delta})\|_{\mathrm{op}}\\ &= 4d\sqrt{nd}\,\|\operatorname{Tr}_d(\boldsymbol{\Delta})\|_{\mathrm{op}}\,d_F(\boldsymbol{S},\boldsymbol{Z})\end{aligned}$$

Now we wrap up the calculations:

$$\begin{aligned}(p-2d)(n^2d-\|\boldsymbol{Z}^\top\boldsymbol{S}\|_F^2) &\le d^2\|\boldsymbol{\Delta}\|_{\mathrm{op}}^2 + 4d\sqrt{nd}\|\operatorname{Tr}_d(\boldsymbol{\Delta})\|_{\mathrm{op}} d_F(\boldsymbol{S},\boldsymbol{Z})\\ &\quad + 2(p-d)\|\boldsymbol{\Delta}\|_{\mathrm{op}}\sqrt{nd}\cdot d_F(\boldsymbol{S},\boldsymbol{Z})\\ &\le d^2\|\boldsymbol{\Delta}\|_{\mathrm{op}}^2 + 2\sqrt{nd}(p-d+2\gamma d)\|\boldsymbol{\Delta}\|_{\mathrm{op}}\cdot d_F(\boldsymbol{S},\boldsymbol{Z})\end{aligned}$$

where $\gamma = \frac{\|\operatorname{Tr}(\boldsymbol{\Delta})\|_{\mathrm{op}}}{\|\boldsymbol{\Delta}\|_{\mathrm{op}}}\vee 1$ is defined in (8). Note that $n^2d - \|\boldsymbol{Z}^\top\boldsymbol{S}\|_F^2 \ge 2^{-1}nd_F^2(\boldsymbol{S},\boldsymbol{Z})$ in (26). Thus for $p > 2d$, we have

$$\frac{nd_F^2(\boldsymbol{S},\boldsymbol{Z})}{2} \le \frac{2\sqrt{nd}(p-d+2\gamma d)\|\boldsymbol{\Delta}\|_{\mathrm{op}}}{p-2d}\cdot d_F(\boldsymbol{S},\boldsymbol{Z}) + \frac{d^2}{p-2d}\|\boldsymbol{\Delta}\|_{\mathrm{op}}^2$$

and equivalently

$$d_F^2(\boldsymbol{S},\boldsymbol{Z}) \le \frac{4(p-d+2\gamma d)}{p-2d}\cdot\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}}\cdot d_F(\boldsymbol{S},\boldsymbol{Z}) + \frac{2d}{p-2d}\cdot\frac{d}{n}\|\boldsymbol{\Delta}\|_{\mathrm{op}}^2.$$

As a result, we have

$$\begin{aligned}d_F(\boldsymbol{S},\boldsymbol{Z}) &\le \left(\frac{2(p-d+2\gamma d)}{p-2d} + \sqrt{4\left(\frac{p-d+2\gamma d}{p-2d}\right)^2 + \frac{2d}{p-2d}}\right)\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}}\\ &\le \frac{(2+\sqrt{5})(p-d+2\gamma d)}{p-2d}\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}} \le \frac{(2+\sqrt{5})(p+d)\gamma}{p-2d}\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}}\end{aligned}$$

where $(p-d+2\gamma d)^2 \ge (p+d)^2 \ge 2d(p-2d)$ holds for $p \ge 2d+1$ and $\gamma \ge 1$. □

5.7 Proof of Theorem 2 and 4

Proposition 2 implies that it suffices to prove

$$n \ge \frac{3\delta^2 d}{2n}\|\boldsymbol{\Delta}\|_{\mathrm{op}}^2 + \delta\sqrt{\frac{d}{n}}\|\boldsymbol{\Delta}\|_{\mathrm{op}}^2 + \max_{1\le i\le n}\left\|\boldsymbol{\Delta}_i^\top\boldsymbol{Z}\right\|_{\mathrm{op}} + \|\boldsymbol{\Delta}\|_{\mathrm{op}} \tag{28}$$

where $\max_{1\le i\le n}\|\boldsymbol{\Delta}_i\|_{\mathrm{op}} \le \|\boldsymbol{\Delta}\|_{\mathrm{op}}$. Now we will estimate $\|\boldsymbol{\Delta}\|_{\mathrm{op}}$ and $\max_{1\le i\le n}\|\boldsymbol{\Delta}_i^\top \boldsymbol{Z}\|_{\mathrm{op}}$ for $\boldsymbol{\Delta} = \sigma \boldsymbol{W}$ where $\boldsymbol{W}$ is an $nd \times nd$ symmetric Gaussian random matrix.

Proof of Theorem 2 and 4 The proof is straightforward: to show that (28) holds for some δ in both convex and nonconvex cases. If $\boldsymbol{\Delta} = \sigma\boldsymbol{W}$ where $\boldsymbol{W}$ is a Gaussian random matrix, it holds that

$$\|\boldsymbol{\Delta}\|_{\mathrm{op}} \le 3\sigma\sqrt{nd}$$

with high probability at least $1 - e^{-nd/2}$ according to [7, Proposition 3.3]. For $\boldsymbol{\Delta}_i^\top\boldsymbol{Z}$, we have

$$\boldsymbol{\Delta}_i^\top \boldsymbol{Z} = \sigma\sum_{j\neq i}\boldsymbol{W}_{ij} \in \mathbb{R}^{d\times d}, \quad (\boldsymbol{\Delta}_i^\top \boldsymbol{Z})_{k\ell} \overset{\text{i.i.d.}}{\sim} \sigma\mathcal{N}(0, n-1), \quad 1 \le k, \ell \le d.$$

Theorem 4.4.5 in [55] implies that the Gaussian matrix $\boldsymbol{\Delta}_i^\top\boldsymbol{Z}$ is bounded by

$$\left\|\boldsymbol{\Delta}_i^\top \boldsymbol{Z}\right\|_{\mathrm{op}} \le C_2\sigma\sqrt{n}\left(\sqrt{d} + \sqrt{2\log n}\right)$$

with probability at least $1 - 2n^{-2}$. By taking the union bound over all $1 \le i \le n$, we have

$$\max_{1\le i\le n}\|\boldsymbol{\Delta}_i^\top \boldsymbol{Z}\|_{\mathrm{op}} \le C_2\sigma\sqrt{n}\left(\sqrt{d} + \sqrt{2\log n}\right)$$

with probability at least $1 - 2n^{-1}$. In the convex relaxation, we have $\delta = 4$. Then the right hand of (28) is bounded by

$$C_1\left(\left(\frac{3\delta^2 d}{2n} + \delta\sqrt{\frac{d}{n}}\right) 9\sigma^2 nd + C_2\sigma\sqrt{n}\left(\sqrt{d} + \sqrt{2\log n}\right) + 3\sigma\sqrt{nd}\right)$$

where $\delta = 4$. The leading term is of order $\sigma^2\sqrt{n}d^{3/2}$ and thus $\sigma < C_0 n^{1/4} d^{-3/4}$ guarantees the tightness of SDP.

For Burer-Monteiro approach, it suffices to estimate γ in (8). The partial trace $\mathrm{Tr}_d(\boldsymbol{\Delta})$ is essentially equal to $\sigma\sqrt{d}\boldsymbol{W}_{GOE,n}$, which implies

$$\|\mathrm{Tr}_d(\boldsymbol{\Delta})\|_{\mathrm{op}} \le 3\sigma\sqrt{nd}$$

with probability at least $1 - e^{-n/2}$ and

$$\delta\|\boldsymbol{\Delta}\|_{\mathrm{op}} \le \frac{\left(2+\sqrt{5}\right)(p+d)}{(p-2d)}\max\{\|\boldsymbol{\Delta}\|_{\mathrm{op}}, \|\mathrm{Tr}_d(\boldsymbol{\Delta})\|_{\mathrm{op}}\} \lesssim \frac{p+d}{p-2d}\cdot\sigma\sqrt{nd}.$$

where $\delta = (2+\sqrt{5})(p+d)(p-2d)^{-1}\gamma$.

Thus the right hand of (28) is bounded by

$$C_1' \left(\frac{d}{n} \cdot \left(\frac{p+d}{p-2d}\right)^2 \cdot \sigma^2 nd + \sqrt{\frac{d}{n}} \cdot \frac{p+d}{p-2d} \cdot \sigma^2 nd + \sigma\sqrt{n}\left(\sqrt{d}+\sqrt{2\log n}\right) + \sigma\sqrt{nd} \right)$$

for some universal constant C_1'. The leading order term is $\sigma^2(p-2d)^{-1}(p+d)d\sqrt{nd}$ which implies that (28) holds if

$$\sigma^2 < \frac{C_0 n(p-2d)}{d\sqrt{nd}(p+d)} = \frac{C_0 n^{1/2}(p-2d)}{d^{3/2}(p+d)}$$

for some small constant C_0. This means $\sigma^2 < C_0 n^{1/2}(p-2d)d^{-3/2}(p+d)^{-1}$ ensures that the optimization landscape of (BM) is benign. □

References

1. Abbe, E.: Community detection and stochastic block models: recent developments. J. Mach. Learn. Res. **18**(1), 6446–6531 (2017)
2. Abbe, E., Bandeira, A.S., Bracher, A., Singer, A.: Decoding binary node labels from censored edge measurements: Phase transition and efficient recovery. IEEE Trans. Netw. Sci. Eng. **1**(1), 10–22 (2014)
3. Absil, P.-A., Mahony, R., Sepulchre, R.: Optimization Algorithms on Matrix Manifolds. Princeton University Press, New Jersey (2009)
4. Arie-Nachimson, M., Kovalsky, S.Z., Kemelmacher-Shlizerman, I., Singer, A., Basri, R.: Global motion estimation from point matches. In: 2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission, pp. 81–88. IEEE (2012)
5. Baik, J., Arous, G.B., Péché, S., et al.: Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. Ann. Probab. **33**(5), 1643–1697 (2005)
6. Bandeira, A.S.: Random Laplacian matrices and convex relaxations. Found. Comput. Math. **18**(2), 345–379 (2018)
7. Bandeira, A.S., Boumal, N., Singer, A.: Tightness of the maximum likelihood semidefinite relaxation for angular synchronization. Math. Program. **163**(1–2), 145–167 (2017)
8. Bandeira, A.S., Boumal, N., Voroninski, V.: On the low-rank approach for semidefinite programs arising in synchronization and community detection. In: Conference on Learning Theory, pp. 361–382 (2016)
9. Arous, G.B., Mei, S., Montanari, A., Nica, M.: The landscape of the spiked tensor model. Commun. Pure Appl. Math. **72**(11), 2282–2330 (2019)
10. Ben-Tal, A., Nemirovski, A.: Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications. SIAM, Philadelphia (2001)
11. Benaych-Georges, F., Nadakuditi, R.R.: The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. Adv. Math. **227**(1), 494–521 (2011)
12. Boumal, N.: A Riemannian low-rank method for optimization over semidefinite matrices with block-diagonal constraints. arXiv preprint arXiv:1506.00575 (2015)
13. Boumal, N.: Nonconvex phase synchronization. SIAM J. Optim. **26**(4), 2355–2377 (2016)
14. Boumal, N., Voroninski, V., Bandeira, A.: The non-convex Burer-Monteiro approach works on smooth semidefinite programs. In: Advances in Neural Information Processing Systems, pp. 2757–2765 (2016)
15. Boumal, N., Voroninski, V., Bandeira, A.S.: Deterministic guarantees for Burer-Monteiro factorizations of smooth semidefinite programs. Commun. Pure Appl. Math. **73**(3), 581–608 (2020)
16. Burer, S., Monteiro, R.D.: A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. Math. Program. **95**(2), 329–357 (2003)
17. Burer, S., Monteiro, R.D.: Local minima and convergence in low-rank semidefinite programming. Math. Program. **103**(3), 427–444 (2005)

18. Candes, E.J., Li, X., Soltanolkotabi, M.: Phase retrieval via wirtinger flow: theory and algorithms. IEEE Trans. Inf. Theory **61**(4), 1985–2007 (2015)
19. Candès, E.J., Romberg, J., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory **52**(2), 489–509 (2006)
20. Capitaine, M., Donati-Martin, C., Féral, D.: The largest eigenvalues of finite rank deformation of large wigner matrices: convergence and nonuniversality of the fluctuations. Ann. Probab. **37**(1), 1–47 (2009)
21. Chaudhury, K.N., Khoo, Y., Singer, A.: Global registration of multiple point clouds using semidefinite programming. SIAM J. Optim. **25**(1), 468–501 (2015)
22. Chen, Y., Candès, E.J.: The projected power method: an efficient algorithm for joint alignment from pairwise differences. Commun. Pure Appl. Math. **71**(8), 1648–1714 (2018)
23. Dellaert, F., Rosen, D.M., Wu, J., Mahony, R., Carlone, L.: Shonan rotation averaging: global optimality by surfing $\mathrm{SO}(p)^n$. In: European Conference on Computer Vision, pp. 292–308. Springer, New York (2020)
24. Gao, B., Liu, X., Chen, X., Yuan, Y.-X.: A new first-order algorithmic framework for optimization problems with orthogonality constraints. SIAM J. Optim. **28**(1), 302–332 (2018)
25. Ge, R., Jin, C., Zheng, Y.: No spurious local minima in nonconvex low rank problems: A unified geometric analysis. In: Proceedings of the 34th International Conference on Machine Learning, vol. 70, pp. 1233–1242 (2017)
26. Giridhar, A., Kumar, P.R.: Distributed clock synchronization over wireless networks: algorithms and analysis. In: Proceedings of the 45th IEEE Conference on Decision and Control, pp. 4915–4920. IEEE (2006)
27. Goemans, M.X., Williamson, D.P.: Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. J. ACM (JACM) **42**(6), 1115–1145 (1995)
28. Huang, Q.-X., Guibas, L.: Consistent shape maps via semidefinite programming. Comput. Graph. Forum **32**(5), 177–186 (2013)
29. Iwen, M.A., Preskitt, B., Saab, R., Viswanathan, A.: Phase retrieval from local measurements: improved robustness via eigenvector-based angular synchronization. Appl. Comput. Harmon. Anal. **48**(1), 415–444 (2020)
30. Jung, J.H., Chung, H.W., Lee, J.O.: Weak detection in the spiked wigner model with general rank. arXiv preprint arXiv:2001.05676 (2020)
31. Keshavan, R.H., Montanari, A., Oh, S.: Matrix completion from a few entries. IEEE Trans. Inf. Theory **56**(6), 2980–2998 (2010)
32. Lerman, G., Shi, Y.: Robust group synchronization via cycle-edge message passing. Found. Comput. Math. (2021)
33. Ling, S.: Generalized power method for generalized orthogonal Procrustes problem: global convergence and optimization landscape analysis. arXiv preprint arXiv:2106.15493 (2021)
34. Ling, S.: Improved performance guarantees for orthogonal group synchronization via generalized power method. SIAM J. Optim **32**(2), 1018–1048 (2022). https://doi.org/10.1137/20M1389571
35. Ling, S.: Near-optimal performance bounds for orthogonal and permutation group synchronization via spectral methods. Appl. Comput. Harmon. Anal. **60**, 20–52 (2022)
36. Ling, S., Xu, R., Bandeira, A.S.: On the landscape of synchronization networks: a perspective from nonconvex optimization. SIAM J. Optim. **29**(3), 1879–1907 (2019)
37. Liu, H., Yue, M.-C., So, A.M.-C.: On the estimation performance and convergence rate of the generalized power method for phase synchronization. SIAM J. Optim. **27**(4), 2426–2446 (2017)
38. Liu, H., Yue, M.-C., So, A.M.-C.: A unified approach to synchronization problems over subgroups of the orthogonal group. arXiv preprint arXiv:2009.07514 (2020)
39. Markdahl, J., Thunberg, J., Goncalves, J.: Almost global consensus on the n-sphere. IEEE Trans. Autom. Control **63**(6), 1664–1675 (2017)
40. Markdahl, J., Thunberg, J., Goncalves, J.: High-dimensional Kuramoto models on Stiefel manifolds synchronize complex networks almost globally. Automatica **113**, 108736 (2020)
41. Mei, S., Misiakiewicz, T., Montanari, A., Oliveira, R.I.: Solving SDPs for synchronization and MaxCut problems via the Grothendieck inequality. In: Conference on Learning Theory, pp. 1476–1515 (2017)
42. Nesterov, Y.: Introductory Lectures on Convex Optimization: A Basic Course, vol. 87. Springer, New York (2013)
43. Pachauri, D., Kondor, R., Singh, V.: Solving the multi-way matching problem by permutation synchronization. In: Advances in Neural Information Processing Systems, pp. 1860–1868 (2013)

44. Pataki, G.: On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. Math. Oper. Res. **23**(2), 339–358 (1998)
45. Perry, A., Wein, A.S., Bandeira, A.S., Moitra, A.: Optimality and sub-optimality of PCA I: spiked random matrix models. Ann. Stat. **46**(5), 2416–2451 (2018)
46. Recht, B., Fazel, M., Parrilo, P.A.: Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. SIAM Rev. **52**(3), 471–501 (2010)
47. Rosen, D.M., Carlone, L., Bandeira, A.S., Leonard, J.J.: Se-sync: a certifiably correct algorithm for synchronization over the special Euclidean group. Int. J. Robot. Res. **38**(2–3), 95–125 (2019)
48. Singer, A.: Angular synchronization by eigenvectors and semidefinite programming. Appl. Comput. Harmon. Anal. **30**(1), 20–36 (2011)
49. Singer, A., et al.: Mathematics for cryo-electron microscopy. Proc. Int. Congr. Math. (ICM) **3**, 3981–4000 (2018)
50. Singer, A., Shkolnisky, Y.: Three-dimensional structure determination from common lines in cryo-em by eigenvectors and semidefinite programming. SIAM J. Imaging Sci. **4**(2), 543–572 (2011)
51. Sun, D., Toh, K.-C., Yuan, Y., Zhao, X.-Y.: SDPNAL+: a Matlab software for semidefinite programming with bound constraints (version 1.0). Optim. Methods Softw. **35**(1), 87–115 (2020)
52. Sun, J., Qu, Q., Wright, J.: Complete dictionary recovery over the sphere i: overview and the geometric picture. IEEE Trans. Inf. Theory **63**(2), 853–884 (2016)
53. Sun, J., Qu, Q., Wright, J.: A geometric analysis of phase retrieval. Found. Comput. Math. **18**(5), 1131–1198 (2018)
54. Tütüncü, R.H., Toh, K.-C., Todd, M.J.: Solving semidefinite-quadratic-linear programs using SDPT3. Math. Program. **95**(2), 189–217 (2003)
55. Vershynin, R.: High-Dimensional Probability: An Introduction with Applications in Data Science, vol. 47. Cambridge University Press, Cambridge (2018)
56. Waldspurger, I., Waters, A.: Rank optimality for the Burer-Monteiro factorization. SIAM J. Optim. **30**(3), 2577–2602 (2020)
57. Wang, L., Singer, A.: Exact and stable recovery of rotations for robust synchronization. Inf. Inference J. IMA **2**(2), 145–193 (2013)
58. Wen, Z., Yin, W.: A feasible method for optimization with orthogonality constraints. Math. Program. **142**(1–2), 397–434 (2013)
59. Xu, R.: On the Landscape of phase synchronization. Master's Thesis. New York University (2019)
60. Yang, L., Sun, D., Toh, K.-C.: SDPNAL+: a majorized semismooth Newton-CG augmented Lagrangian method for semidefinite programming with nonnegative constraints. Math. Program. Comput. **7**(3), 331–366 (2015)
61. Yurtsever, A., Tropp, J.A., Fercoq, O., Udell, M., Cevher, V.: Scalable semidefinite programming. SIAM J. Math. Data Sci. **3**(1), 171–200 (2021)
62. Zhang, T.: Tightness of the semidefinite relaxation for orthogonal trace-sum maximization. arXiv preprint arXiv:1911.08700 (2019)
63. Zhong, Y., Boumal, N.: Near-optimal bounds for phase synchronization. SIAM J. Optim. **28**(2), 989–1016 (2018)