A Unified Complexity Metric for Nonconvex Matrix Completion and Matrix Sensing in the Rank-one Case

Haixiang Zhang · Baturalp Yalcin · Javad Lavaei · Somayeh Sojoudi

Abstract In this work, we develop a new complexity metric for an important class of low-rank matrix optimization problems, where the metric aims to quantify the complexity of the nonconvex optimization landscape of each problem and the success of local search methods in solving the problem. The existing literature has focused on two complexity measures. The RIP constant is commonly used to characterize the complexity of matrix sensing problems. On the other hand, the sampling rate and the incoherence are used when analyzing matrix completion problems. The proposed complexity metric has the potential to unify these two notions and also applies to a much larger class of problems. To mathematically study the properties of this metric, we focus on the rank-1 generalized matrix completion problem and illustrate the usefulness of the new complexity metric from three aspects. First, we show that instances with the RIP condition have a small complexity. Similarly, if the instance obeys the

Corresponding Author: Javad Lavaei

Affiliation: Department of Industrial Engineering and Operations Research, University of California, Berkeley

E-mail: lavaei@berkeley.edu

Haixiang Zhang
Department of Mathematics, University of California, Berkeley, CA
E-mail: haixiang_zhang@berkeley.edu

Baturalp Yalcin
Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA
E-mail: baturalp_yalcin@berkeley.edu

Javad Lavaei
Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA
E-mail: lavaei@berkeley.edu

Somayeh Sojoudi
Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA
E-mail: sojoudi@berkeley.edu

Bernoulli sampling model, the complexity metric will take a small value with high probability. Moreover, for a one-parameter class of instances, the complexity metric shows consistent behavior to the first two scenarios. Furthermore, we establish theoretical results to provide sufficient conditions and necessary conditions on the existence of spurious solutions in terms of the proposed complexity metric. This contrasts with the RIP and incoherence notions that fail to provide any necessary condition.

Keywords Matrix sensing · Matrix completion · Complexity metric · Nonconvex optimization · Global convergence

Mathematics Subject Classification (2020) 05C90 · 65F55 · 90C26

1 Introduction

A variety of modern signal processing and machine learning applications require solving optimization problems that involve a low-rank matrix variable. More specifically, given measurements to some unknown ground truth matrix $M^* \in \mathbb{R}^{n \times n}$ of rank $r \ll n$, the *low-rank matrix optimization* problem can be formulated as

$$(1.1) \qquad \min_{M \in \mathbb{R}^{n \times n}} f(M; M^*) \quad \text{s.t.} \quad M \succeq 0, \quad \text{rank}(M) \leq r,$$

where $f(\cdot; M^*)$ is the loss function that penalizes the mismatch between the input matrix and M^*. The goal is to recover the matrix M^* via (1.1). Examples of this problem include matrix sensing [42,59,57], matrix completion [14,15,26], phase retrieval [12,45,20] and robust principle component analysis [11,25]; see the review papers [18,23] for more applications. To deal with the nonconvex rank constraint, there have been several works on the convex relaxations of problem (1.1). More concretely, one may replace the rank constraint with a nuclear norm regularizer [14,42,15,11]. The convex relaxation approach is proven to achieve the optimal sampling complexity for various statistical models. However, the convex relaxation approach needs to update a matrix variable in each iteration, which relies on the Singular Value Decomposition (SVD) of the matrix variable. This will lead to an $O(n^3)$ computational complexity in each iteration and an $O(n^2)$ space complexity, which are prohibitively high for large-scale problems; see the numerical comparison in [60].

To improve the computational efficiency, an alternative approach was proposed by Burer and Monteiro [10], which is named as the Burer-Monteiro factorization approach. The factorization approach is based on the fact that the mapping $U \mapsto UU^T$ is surjective onto the manifold of positive semi-definite matrices of rank at most r, where $U \in \mathbb{R}^{n \times r}$. Therefore, problem (1.1) is equivalent to

$$(1.2) \qquad \min_{U \in \mathbb{R}^{n \times r}} f(UU^T; M^*),$$

which is an unconstrained nonconvex problem. A major difficulty about nonconvex optimization problems is the existence of spurious local minima[1]. In general, common local search methods are only able to guarantee a point approximately satisfying

---

[1] A point U^0 is called a spurious local minimum if it is a local minimum of problem (1.2) and $U^0(U^0)^T \neq M^*$.

the first-order and the second-order necessary optimality conditions. Therefore, local search methods with a random initialization will likely be stuck at spurious local minima and unable to converge to the global solution. However, despite the aforementioned issue of nonconvex optimization problems, simple iterative algorithms such as gradient descent and alternating minimization have achieved empirical success in a wide range of applications. In recent years, substantial progress has been made on the theoretical understandings of these algorithms, which generally focused on proving the in-existence of spurious local minima. For example, the alternating minimization algorithm was first studied in [32, 39, 40]. The (stochastic) gradient descent algorithm, which is in general easier to implement than the alternating minimization algorithm, was analyzed in [12, 50, 56, 20, 18]. Besides algorithmic analysis, a critical geometric property named the strict-saddle property [45] was established in [26, 45, 61, 57], which can guarantee the polynomial-time global convergence of various saddle-escaping algorithms [16, 34, 5].

To characterize the behavior of local search methods for problem (1.2), several complexity metrics were proposed. A small complexity metric implies that the landscape of problem (1.2) is benign and thus, local search methods with random initialization converge to global solutions with high probability. Otherwise, if the complexity metric takes a large value, problem (1.2) may have spurious local minima, which will imply the failure of most local search methods. However, the complexity metrics for problem (1.2) are designed separately for different applications. As a result, several different metrics were proposed to characterize the optimization complexity of problem (1.2). For example, in the context of matrix sensing problems, the following Restrict Isometry Property (RIP) is usually assumed:

Definition 1.1 ([42,61]) Given natural numbers r and s, the function $f(\cdot; M^*)$ is said to satisfy the **Restricted Isometry Property** (RIP) of rank $(2r, 2s)$ for a constant $\delta \in [0, 1)$, denoted as δ-RIP$_{2r,2s}$, if

$$(1.3) \qquad (1 - \delta)\|K\|_F^2 \leq \left[\nabla^2 f(M; M^*)\right](K, K) \leq (1 + \delta)\|K\|_F^2$$

holds for all matrices $M, K \in \mathbb{R}^{n \times n}$ such that $\text{rank}(M) \leq 2r, \text{rank}(K) \leq 2s$, where $\left[\nabla^2 f(M; M^*)\right](\cdot, \cdot)$ is the curvature of the Hessian at point M.

One important class of matrix sensing problems is the *linear matrix sensing problem*, which is induced by linear measurements of the ground truth matrix M^*. If the ℓ_2-loss is used, the linear matrix sensing problem can be formulated as

$$(1.4) \qquad \min_{U \in \mathbb{R}^{n \times r}} \frac{1}{m} \sum_{i=1}^{m} \langle A_i, UU^T - M^* \rangle^2,$$

where $m \in \mathbb{N}$ is the number of measurements modeled by the known measurement matrices $A_i \in \mathbb{R}^{n \times n}$ for all $i \in [m]$. In the special case when each matrix A_i is an independently identically distributed Gaussian random matrix, the δ-RIP$_{2r,2s}$ condition holds with high probability if $m = O(nr\delta^{-2})$ [13]. The RIP constant δ plays a critical role in measuring the optimization complexity of problem (1.2). In [7], the authors showed that the strict-saddle property holds for problem (1.2) if the δ-RIP$_{2r,2r}$ condition holds with $\delta < 1/2$ and the ground truth matrix satisfies $\text{rank}(M^*) = r$.

On the other hand, counterexamples have been constructed in [59,57] to illustrate that the strict-saddle property can fail under the δ-RIP$_{2r,2r}$ condition with $\delta \geq 1/2$.

Despite these strong theoretical results under the RIP assumption, there exists a large number of applications that do not satisfy the RIP condition. One of those applications without the RIP condition is the *matrix completion problem*. Given a set of indices $\Omega \subset [n] \times [n]$, the matrix completion problem aims at recovering the low-rank matrix M^* from the available entries M_{ij}^* for $(i,j) \in \Omega$. With the least squares loss function, the matrix completion problem can be formulated as

$$(1.5) \qquad \min_{U \in \mathbb{R}^{n \times r}} \sum_{(i,j) \in \Omega} \left[(UU^T)_{ij} - M_{ij}^* \right]^2.$$

The matrix completion problem (1.5) is a special case of the matrix sensing problem (1.4), where each measurement matrix A_i has exactly one nonzero entry. However, the RIP$_{2r,2r}$ condition does not hold for problem (1.5) unless all entries of M^* are observed, namely, when $\Omega = [n] \times [n]$. As an alternative to the RIP condition, the optimization complexity of problem (1.5) is closely related to the incoherence of M^*.

Definition 1.2 ([14]) Given a constant $\mu \in [1, n]$, the ground truth matrix M^* is said to be μ-**incoherent** if

$$(1.6) \qquad \|e_i^T V^*\|_F \leq \sqrt{\mu r/n}, \quad \forall i \in [n],$$

where $V^* \Lambda^* (V^*)^T$ is the truncated SVD of M^* and e_i is the i-th standard basis of $\mathbb{R}^n$.

Intuitively, if the ground truth M^* is highly sparse, it is likely that only zero entries of M^* are observed and there is no chance to learn the other entries of the matrix M^*. A relatively small incoherence of M^* avoids this extreme case. The most popular statistical model of the measurements for problem (1.5) is the Bernoulli model, where each entry of M^* is observed independently with probability $p \in (0, 1]$. Assuming the Bernoulli model, the scaled gradient descent algorithm with a spectral initialization [47] converges linearly given the condition $p \geq O(\mu r^2 \kappa^2 \max(\mu \kappa^2, \log n)/n)$, where $\kappa := \sigma_1(M^*)/\sigma_r(M^*)$ is the condition number of M^*. In addition, the global convergence was established in [26] through the strict-saddle property under the assumption that $p \geq O(\mu^4 r^6 \kappa^6 \log n/n)$. We note that the dependence on the condition number κ may be unnecessary as shown in [29] and that the condition number is equal to 1 in the rank-1 case. On the other hand, the information-theoretical lower bound in [14] shows that $p \geq \Theta(\mu r \log(n/\delta)/n)$ is necessary for the exact completion with probability at least $1 - \delta$. Therefore, the complexity of problem (1.5) is mainly measured by the incoherence of M^*.

The main issue with the notions of RIP and incoherence is that they require stringent conditions to guarantee the success of local search methods for recovering M^*. Whenever these conditions are violated, local search methods may still work successfully, which questions whether these customized notions designed for special cases of the problem truly capture the complexity of the problem in general. Hence, it is natural to ask:

Does there exist a complexity metric with two properties: (i) it unifies the metrics for problems (1.4) and (1.5), namely, the RIP constant δ and the incoherence μ, (ii) even when the RIP and incoherence conditions are violated, it still quantifies the optimization complexity of the problem in the sense that the smaller the value of this metric is, the higher the success of local search methods with random initialization is in finding the ground truth M^*?

In this work, we provide a partial answer to the question by developing a powerful complexity metric. To analyze the usefulness of this new metric, we focus on the generalized rank-1 matrix completion problem

$$(1.7) \qquad \min_{u \in \mathbb{R}^n} \sum\nolimits_{i,j \in [n]} C_{ij}(u_i u_j - M_{ij}^*)^2,$$

where the ground truth M^* is symmetric and has rank at most 1. The weights are $C_{ij} \geq 0$ for all $i, j \in [n]$. Without loss of generality, we can assume that the matrix $C := (C_{ij})_{i,j \in [n]}$ is symmetric since otherwise one can replace C with $(C + C^T)/2$, which will not change the optimization landscape. We use $\mathcal{MC}(C, u^*)$ to denote the instance of problem (1.7) with the weight matrix C and the ground truth $M^* = u^*(u^*)^T$, for all $C \in \mathbb{R}^{n \times n}$ and $u^* \in \mathbb{R}^n$. The matrix completion problem (1.5) is a special case of the generalized matrix completion problem (1.7), where $C_{ij} = 1$ if $(i, j) \in \Omega$ and $C_{ij} = 0$ otherwise. Moreover, problem (1.7) is a special case of the matrix sensing problem (1.4), where each measurement only captures one entry of M^*. However, in Section 3.3, we show that there exists an instance of problem (1.7) that satisfies the $1/2$-RIP$_{2,2}$ condition but has spurious local minima. This counterexample implies that the optimal RIP bound in [59, 57] still holds for problem (1.7) and thus, problem (1.7) still contains difficult instances of the matrix sensing problem. Indeed, we show in Section 3.1 that some of the results to be developed for problem (1.7) can be extended to the general matrix sensing problem (1.4).

Now, we provide an intuition into the design of our complexity metric for problem (1.7). For a given problem instance of (1.7), if there exist global solutions u^1, u^2 such that $u^1(u^1)^T \neq u^2(u^2)^T$, it is impossible to decide which global solution corresponds to M^* from the observations. Intuitively, no matter what optimization algorithm we choose and how much computational effort is exerted, there is a chance that we could not recover M^* by solving problem (1.7). This observation motivates us to define the complexity metric to be the inverse of the infimum of the distance between any given instance and the set of instances with multiple global solutions. Since problem (1.7) is parameterized by the weight matrix C and the global solution M^*, we are able to define the metric through norms in Euclidean spaces and their Cartesian products. In addition, in the rank-1 case, the (random) graph theory serves as an important tool in characterizing the solvability of problem (1.7). These two advantages enable a more thorough analysis of the new complexity metric. The formal definition of the metric is provided in Section 2. In this work, we exhibit several pieces of evidence to show that the proposed metric can serve as an alternative to the RIP constant and the incoherence, which are summarized below:

1. For problem instances that satisfy the δ-RIP$_{2,2}$ condition, we provide an upper bound on the complexity metric. The upper bound is tightened with extra information about the incoherence of M^*. The complexity metric is also generalized

to problem (1.2) and a slightly weaker upper bound is derived. Similarly, for matrix completion problems obeying the Bernoulli model, an upper bound on the complexity metric in terms of the incoherence of M^* is derived.

2. We then construct a class of parameterized instances of problem (1.7). A lower bound on the complexity metric is developed to prove that instances whose complexity metric is larger than the lower bound have an exponential number of spurious local minima. In addition, an upper bound that is consistent with the existing results for the matrix sensing and matrix completion problems are established to guarantee the in-existence of spurious local minima if the complexity metric is below this bound.

3. We prove the existence of a non-trivial upper bound on the complexity metric. For all problem instances whose complexity metric is below this upper bound, problem (1.7) has no spurious local minima and M^* can be successfully found via local search methods with random initialization.

4. Under a standard bounded-away-from-zero assumption, we show that all instances with a larger complexity metric will possess spurious local minima.

Based on the aforementioned results, we make some key conjectures and discuss the potential extensions of the proposed metric to more general cases of the low-rank matrix optimization problem (1.1).

1.1 Related works

Following the famous *Netflix prize*, the theoretical analysis of problem (1.1) has attracted a lot of attention in recent years; see the review papers [19,23]. Early attempts mainly focused on the construction of convex relaxations to rank-constrained problems [14,15,42,11], where the RIP condition and the incoherence were introduced. Although the convex relaxation is usually guaranteed to recover the exact ground truth with almost the optimal sample complexity, the associated algorithms operate in the space of matrix variables and, thus, are computationally inefficient for large-scale problems [60]. Similar issues are observed for algorithms based on the Singular Value Projection [31] and Riemannian optimization algorithms [53,54,30,2,37]. The analysis of the convex relaxation approach in the noisy case is recently conducted by bridging the convex and the nonconvex approaches [21,22].

To deal with the difficulties in solving large-scale problems, an efficient alternative model (1.2) using the Burer-Monteiro factorization is considered. Despite the nonconvexity, a growing number of works demonstrated that problem (1.2) has benign landscapes and, therefore, is amenable for efficient optimization. Theoretical analysis stems from the alternating minimization method [32,39,28,29,40,1]. The alternating minimization method has the advantage that the number of iterations has only logarithmic dependence on the condition number of the ground truth [29]. More recently, this advantage is also achieved by the scaled (sub)gradient descent algorithm [47,48,49,58].

The gradient descent algorithm has also gained a significant attention due to its simplicity in implementation. In general, there are two ways to apply the gradient

descent algorithm. First, the gradient descent algorithm can serve as the local refinement method after a suitable initialization [12,50,46,56,4,18]. On the other hand, the gradient descent algorithm is proved to converge globally for the phase retrieval problem [20]. More generally, under the strict-saddle property, a number of saddle-escaping algorithms [34,16,5] converge to the global solution in polynomial time; see e.g., [44,27,26,61,45,59,17,57,6,7,38]. Moreover, the gradient descent algorithm is proved to have the implicit regularization phenomenon in the over-parameterization case [36,24,43].

Finally, the Burer-Monteiro factorization approach has also been studied in the context of finding low-rank solutions to semi-definite programs [8,41,9,52].

1.2 Notation and organization

The number of elements in a finite set $\mathcal{S}$ is denoted as $|\mathcal{S}|$. We use $\overline{\mathcal{S}}$ to denote the closure of a set $\mathcal{S} \subset \mathbb{R}^n$. The index set $\{1, \ldots, n\}$ is denoted as $[n]$ for all $n \in \mathbb{N}$. The entry-wise ℓ_1-norm and the Frobenius norm of a matrix M are denoted as $\|M\|_1$ and $\|M\|_F$, respectively. The unit sphere of matrices with non-negative entries denoted as $\mathbb{S}^{n^2-1}_{+,1}$ is the set of all symmetric matrices $X \in \mathbb{R}^{n \times n}$ such that $\|X\|_1 = 1$ and $X_{ij} \geq 0$ for all $i, j \in [n]$. Similarly, the unit sphere of vectors $\mathbb{S}^{n-1}_1$ is the set of all vectors $x \in \mathbb{R}^n$ such that $\|x\|_1 = 1$. For every symmetric matrix $M \in \mathbb{R}^{n \times n}$, the minimum eigenvalue is denoted as $\lambda_{min}(M)$. The n-by-n identity matrix is denoted as $\mathcal{I}_n$. The notation $M \succeq 0$ means that the matrix M is symmetric and positive semi-definite. The sub-matrix $R_{i:j,k:\ell}$ consists of the i-th to the j-th rows and the k-th to the ℓ-th columns of matrix R. For every vector $x \in \mathbb{R}^n$, the sets of indices corresponding to zero and nonzero components of x are denoted as $\mathcal{I}_0(x)$ and $\mathcal{I}_1(x)$, respectively. For every instance $\mathcal{MC}(C, u^*)$, we use $\mathbb{G}(C, u^*) = [\mathbb{V}(C, u^*), \mathbb{E}(C, u^*), \mathbb{W}(C, u^*)]$ to denote the associated weighted graph, which is defined in Section 2. The unweighted undirected graph $\mathbb{G}$ with node set $\mathbb{V}$ and edge set $\mathbb{E}$ is denoted as $\mathbb{G} = (\mathbb{V}, \mathbb{E})$. The objective function of an instance $\mathcal{MC}(C, u^*)$ is shown as $g(u; C, u^*) := \sum_{i,j \in [n]} C_{ij}(u_i u_j - u_i^* u_j^*)^2$. The action of the Hessian $\nabla^2 g(M; C, u^*)$ on any two matrices K and L is given by $[\nabla^2 g(M; C, u^*)](K, L) := \sum_{i,j,k,\ell} [\nabla^2 g(M; C, u^*)]_{i,j,k,\ell} K_{ij} L_{k,\ell}$. The notations $a_n = O(b_n)$ and $a_n = \Theta(b_n)$ mean that there exist constants $c_1, c_2 > 0$ such that $a_n \leq c_2 b_n$ and $c_1 b_n \leq a_n \leq c_2 b_n$ hold for all $n \in \mathbb{Z}$, respectively.

In the remainder of this paper, we first define the proposed complexity metric and derive basic properties of the metric in Section 2. In Section 3, we analyze this metric on well-studied problems, including the matrix sensing problem and the matrix completion problem. Section 4 is devoted to the theoretical guarantees provided by the new complexity metric. Finally, we conclude the paper in Section 5. Some of the proofs are provided in the appendix.

2 New complexity metric and basic properties

In this section, we provide the formal definition of the new complexity and investigate the properties of the proposed metric. Before proceeding to the definitions, we note that the problem (1.7) is "scale-free" in the sense that the instance $\mathcal{MC}(\eta_1 C, \eta_2 u^*)$ has the same landscape as $\mathcal{MC}(C, u^*)$ up to a scaling, where $C \in \mathbb{R}^{n \times n}$, $u^* \in \mathbb{R}^n$ and $\eta_1, \eta_2 > 0$ are constants. Therefore, we may normalize the parameters C and u^* without loss of generality, as follows:

Assumption 2.1 *Assume that* $C \in \mathbb{S}^{n^2-1}_{+,1}$ *and* $u^* \in \mathbb{S}^{n-1}_1$, *i.e.,* $\|C\|_1 = \|u^*\|_1 = 1$.

The above assumption excludes the degenerate cases when $C = 0$ or $M^* = 0$. If $C = 0$, the objective function is always 0 and it is impossible to recover the ground truth. For the case when $M^* = 0$, we can prove that either $u = 0$ is the only stationary point or the instance $\mathcal{MC}(C, 0)$ has multiple different global solutions. In the first situation, the results in [35] imply that randomly initialized gradient descent algorithm will converge to 0 with probability 1. In the second situation, the instance is information-theoretically unsolvable. We provide a more detailed analysis in the appendix and assume that Assumption 2.1 holds in the remainder of the paper.

The definition of the complexity metric is closely related to the set of instances with multiple "essentially different" global solutions. More specifically, the set of degenerate instances is defined as

$$\mathcal{D} := \{(C, u^*) \mid C \in \mathbb{S}^{n^2-1}_{+,1}, u^* \in \mathbb{S}^{n-1}_1,$$
$$\exists u \in \mathbb{R}^n \quad \text{s.t.} \quad g(u; C, u^*) = 0, \ uu^T \neq u^*(u^*)^T\}.$$

Since there exist multiple global solutions to problem (1.7) if $(C, u^*) \in \mathcal{D}$, it is information-theoretically impossible to find the ground truth for any instance in $\mathcal{D}$. Intuitively, we say that the *optimization complexity* of all instances in $\mathcal{D}$ is infinity. Motivated by the above observation, we introduce the new complexity metric.

Definition 2.1 (Complexity Metric) Given arbitrary parameters $C \in \mathbb{S}^{n^2-1}_{+,1}$, $u^* \in \mathbb{S}^{n-1}_1$ and $\alpha \in [0, 1]$, the complexity of the instance $\mathcal{MC}(C, u^*)$ is defined as

$$(2.1) \qquad \mathbb{D}_\alpha(C, u^*) := \left[\inf_{(\tilde{C}, \tilde{u}^*) \in \mathcal{D}} \alpha \|C - \tilde{C}\|_1 + (1 - \alpha)\|u^* - \tilde{u}^*\|_1 \right]^{-1}.$$

Since the set $\mathcal{D}$ is bounded, the infimum in the definition exists and is finite. The term inside the inverse operation can be viewed as a weighted distance between the point (C, u^*) and the set $\mathcal{D}$. In addition, we take the convention that $1/0 = +\infty$ and thus, $\mathbb{D}_\alpha(C, u^*) = +\infty$ for all $(C, u^*) \in \mathcal{D}$. In this work, we choose the entry-wise ℓ_1-norm in (2.1) for the simplicity of calculations. We believe that similar theory can still be derived for other choices of the norm.

2.1 Basic properties of the new complexity metric

We first provide a more concrete definition of the set $\mathcal{D}$. In the rank-1 case, we are able to exactly describe the set $\mathcal{D}$ using graph-theoretic notations. We first introduce the associated graphs of any instance of the problem. Given an instance $\mathcal{MC}(C, u^*)$, the weighted graph $\mathbb{G}(C, u^*) = [\mathbb{V}(C, u^*), \mathbb{E}(C, u^*), \mathbb{W}(C, u^*)]$ is defined by

$$\mathbb{V}(C, u^*) := [n], \quad \mathbb{E}(C, u^*) := \{\{i, j\} \mid C_{ij} > 0, i, j \in [n]\},$$
$$[\mathbb{W}(C, u^*)]_{ij} := C_{ij}, \quad \forall i, j \in [n] \quad \text{s.t.} \quad \{i, j\} \in \mathbb{E}(C, u^*).$$

To include the information of u^*, we define

$$\mathcal{I}_1(C, u^*) := \{i \in [n] \mid u_i^* \neq 0\}, \quad \mathcal{I}_0(C, u^*) := [n] \backslash \mathcal{I}_1(C, u^*),$$
$$I_{00}(C, u^*) := \{i \in \mathcal{I}_0(C, u^*) \mid \{i, j\} \notin \mathbb{E}(C, u^*), \ \forall j \in \mathcal{I}_1(C, u^*)\}.$$

Intuitively, the sets $\mathcal{I}_1(C, u^*)$ and $\mathcal{I}_0(C, u^*)$ contain the locations of the nonzero and zero components of u^*. The subset $\mathcal{I}_{00}(C, u^*)$ corresponds to indices in $\mathcal{I}_0(C, u^*)$ that are not connected to any index in $\mathcal{I}_1(C, u^*)$. We denote the subgraph of $\mathbb{G}(C, u^*)$ induced by the index set $\mathcal{I}_1(C, u^*)$ as $\mathbb{G}_1(C, u^*) = [\mathcal{I}_1(C, u^*), \mathbb{E}_1(C, u^*), \mathbb{W}_1(C, u^*)]$, where $\mathbb{E}_1(C, u^*)$ and $\mathbb{W}_1(C, u^*)$ are the edge set and weight set of this subgraph. The following theorem provides an equivalent definition of $\mathcal{D}$ in terms of $\mathcal{I}_{00}(C, u^*)$ and $\mathbb{G}_1(C, u^*)$.

Theorem 2.2 *Given $C \in \mathbb{S}_{+,1}^{n^2-1}$ and $u^* \in \mathbb{S}_1^{n-1}$, it holds that $(C, u^*) \notin \mathcal{D}$ if and only if*

1. $\mathbb{G}_1(C, u^*)$ is connected and non-bipartite;
2. $\{i, i\} \in \mathbb{E}(C, u^*)$ for all $i \in \mathcal{I}_{00}(C, u^*)$.

Proof We first construct counterexamples for the necessity part and then prove the uniqueness of the global minimum (up to a sign flip) for the sufficiency part. For the notational simplicity, we fix the point (C, u^*) and omit them in the notations.

Necessity. In this part, our goal is to construct a solution $u \in \mathbb{R}^n$ such that

$$u_i u_j = u_i^* u_j^*, \quad \forall \{i, j\} \in \mathbb{E}; \quad u u^T \neq u^* (u^*)^T.$$

We denote $M^* := u^* (u^*)^T$ and analyze three different cases below.

Case I. First, we consider the case when $\mathbb{G}_1$ is disconnected, which means that there exist two non-empty subsets $\mathcal{I}$ and $\mathcal{J}$ such that

$$\mathcal{I} \cup \mathcal{J} = \mathcal{I}_1, \quad \mathcal{I} \cap \mathcal{J} = \emptyset; \quad \{i, j\} \notin \mathbb{E}_1, \quad \forall i \in \mathcal{I}, \ \forall j \in \mathcal{J}.$$

We define the vector $u \in \mathbb{R}^n$ as

$$u_i := 0, \quad \forall i \in \mathcal{I}_0; \quad u_i = u_i^*, \quad \forall i \in \mathcal{I}; \quad u_i = -u_i^*, \quad \forall i \in \mathcal{J}.$$

The above definition leads to

$$u_i u_j = \begin{cases} -M_{ij}^* & \text{if } i \in \mathcal{I} \text{ and } j \in \mathcal{J} \\ M_{ij}^* & \text{otherwise.} \end{cases}$$

Since $u_i^* \neq 0$ for all $i \in \mathcal{I}_1$, it follows that $u_i u_j = -M_{ij}^* \neq M_{ij}^*$ for all $\{i, j\}$ such that $i \in \mathcal{I}$ and $j \in \mathcal{J}$.

Case II. Next, we consider the case when $\mathbb{G}_1$ is bipartite, which means that there exist two non-empty subsets $\mathcal{I}$ and $\mathcal{J}$ such that

$$\mathcal{I} \cup \mathcal{J} = \mathcal{I}_1, \quad \mathcal{I} \cap \mathcal{J} = \emptyset; \quad \{i, j\} \notin \mathbb{E}_1, \quad \forall i, j \in \mathcal{I}_1 \quad \text{s.t.} \quad i, j \in \mathcal{I} \text{ or } i, j \in \mathcal{J}.$$

In this case, we define the vector $u \in \mathbb{R}^n$ as

$$u_i := 0, \quad \forall i \in \mathcal{I}_0; \quad u_i := u_i^*/2, \quad \forall i \in \mathcal{I}; \quad u_i := 2u_i^*, \quad \forall i \in \mathcal{J}.$$

Now, we have

$$u_i u_j = \begin{cases} M_{ij}^*/4 & \text{if } i, j \in \mathcal{I} \\ 4M_{ij}^* & \text{if } i, j \in \mathcal{J} \\ M_{ij}^* & \text{otherwise.} \end{cases}$$

Since $M_{ij}^* \neq 0$ for all $i, j \in \mathcal{J}$, we have that $u_i u_j = 4M_{ij}^* \neq M_{ij}^*$ for all $i, j \in \mathcal{J}$.

Case III. Finally, we check the case when there exists a node $i_0 \in \mathcal{I}_{00}$ such that $\{i_0, i_0\} \notin \mathbb{E}$. In this case, we define the vector $u \in \mathbb{R}^n$ as

$$u_{i_0} := 1, \quad u_i := u_i^*, \quad \forall i \in [n] \backslash \{i_0\}.$$

Now, we have

$$u_{i_0} u_{i_0} = 1 \neq 0 = M_{i_0 i_0}^*, \quad u_i u_j = M_{ij}^*, \quad \forall \{i, j\} \in \mathbb{E}.$$

Combining the above three cases completes the proof of the necessity part.

Sufficiency. We prove that any global solution $u \in \mathbb{R}^n$ to problem (1.7) satisfies $uu^T = M^*$, where $M^* := u^*(u^*)^T$. Since u is a global solution, it follows that

$$u_i u_j = M_{ij}^*, \quad \forall i, j \in [n] \quad \text{s.t.} \quad \{i, j\} \in \mathbb{E}.$$

Since the graph $\mathbb{G}_1$ is non-bipartite, there exists a cycle with an odd number of edges in $\mathbb{G}_1$. We denote the length of the cycle as $2k + 1$, where k is a non-negative integer. Moreover, we denote the edges of the cycle as

$$\{i_0, i_1\}, \{i_1, i_2\}, \ldots, \{i_{2k}, i_0\}.$$

Since $\{i_0, \ldots, i_{2k}\} \subset \mathcal{I}_1$, we know that

$$u_i u_j = M_{ij}^* \neq 0, \quad \forall i, j \in [n] \quad \text{s.t.} \quad \{i, j\} \in \{\{i_\ell, i_{\ell+1}\}, \ell \in \{0, \ldots, 2k\}\},$$

where $i_{2k+1} := i_0$. Hence, we can calculate that

$$u_0^2 = \prod_{\ell=0}^{2k} (u_{i_\ell} u_{i_{\ell+1}})^{(-1)^\ell} = \prod_{\ell=0}^{2k+1} M_{i_\ell i_{\ell+1}}^{(-1)^\ell} = (u_{i_0}^*)^2.$$

Without loss of generality, assume that $u_{i_0} = u_{i_0}^*$ since otherwise we can consider the solution $-u$ if $u_{i_0} = -u_{i_0}^*$. With the value of u_{i_0} correctly recovered, it follows that

$$u_{i_1} = \frac{u_{i_0} u_{i_1}}{u_{i_0}} = \frac{u_{i_0}^* u_{i_1}^*}{u_{i_0}^*} = u_{i_1}^*.$$

Similarly, we can utilize the connectivity of $\mathbb{G}_1$ to iteratively obtain $u_i = u_i^*$ for all $i \in \mathcal{I}_1$.

The remaining part is to show that $u_i = 0$ for all $i \in \mathcal{I}_0$. For every node $i \in \mathcal{I}_0 \backslash \mathcal{I}_{00}$, there exists a node $j \in \mathcal{I}_1$ such that $\{i, j\} \in \mathbb{E}$. This implies that

$$u_j = u_j^* \neq 0, \quad u_i u_j = M_{ij}^* = 0,$$

Hence, it holds that $u_i = 0$. For every node $i \in \mathcal{I}_{00}$, the assumption in the theorem requires that $\{i, i\} \in \mathbb{E}$, which leads to

$$u_i^2 = M_{ii}^* = 0.$$

In this case, we also obtain $u_i = 0$. $\qquad\square$

Since the set $\mathcal{D}$ is bounded, the infimum in the definition (2.1) can be attained by using the closure of $\mathcal{D}$, namely

(2.2) $$\mathbb{D}_\alpha(C, u^*) = \left[\min_{(\tilde{C}, \tilde{u}^*) \in \overline{\mathcal{D}}} \alpha \|C - \tilde{C}\|_1 + (1 - \alpha)\|u^* - \tilde{u}^*\|_1 \right]^{-1}.$$

The alternative definition (2.2) simplifies the verification of parameters that attain the infimum. In addition, with the help of Theorem 2.2, we can exactly characterize the closure $\overline{\mathcal{D}}$, which has a slightly simpler form than $\mathcal{D}$.

Theorem 2.3 *We have the following relation:*

$$\overline{\mathcal{D}} = \{(C, u^*) \mid C \in \mathbb{S}_{+,1}^{n^2-1}, u^* \in \mathbb{S}_1^{n-1}, \mathbb{G}_1(C, u^*) \text{ is disconnected or bipartite}\}$$
$$\cup \{(C, u^*) \mid C \in \mathbb{S}_{+,1}^{n^2-1}, u^* \in \mathbb{S}_1^{n-1}, \mathcal{I}_{00}(C, u^*) \text{ is not empty}\}.$$

Proof We denote the set on the right-hand side as $\mathcal{D}'$. We first prove that

(2.3) $$\overline{\mathcal{D}} \supset \mathcal{D}'.$$

Suppose that $(C, u^*) \in \mathcal{D}'$. If $\mathbb{G}_1(C, u^*)$ is disconnected or bipartite, the instance $\mathcal{MC}(C, u^*)$ already belongs to $\mathcal{D}$ and, therefore, belongs to the closure $\overline{\mathcal{D}}$. We only need to consider the case when $\mathcal{I}_{00}(C, u^*)$ is not empty. For every constant $\epsilon > 0$, we construct a new global solution $\tilde{u}^*$ as follows:

$$\tilde{u}_i^* := \begin{cases} u_i^* + \epsilon & \text{if } i \in \mathcal{I}_{00}(C, u^*) \\ u_i^* & \text{otherwise.} \end{cases}$$

Let $\tilde{M}^* := \tilde{u}^*(\tilde{u}^*)^T$. For the instance $\mathcal{MC}(C, \tilde{u}^*)$, we have

$$\mathcal{I}_1(C, \tilde{u}^*) = \mathcal{I}_1(C, u^*) \cup \mathcal{I}_{00}(C, u^*).$$

By the definition of $\mathcal{I}_{00}(C, u^*)$, the nodes in $\mathcal{I}_1(C, u^*)$ and $\mathcal{I}_{00}(C, u^*)$ are disconnected. Therefore, the new subgraph $\mathbb{G}_1(C, \tilde{u}^*)$ is disconnected and the new instance $\mathcal{MC}(C, \tilde{u}^*)$ belongs to $\mathcal{D}$. By letting $\epsilon \to 0$, it follows that (C, u^*) is a limit point of $\mathcal{D}$ and belongs to $\overline{\mathcal{D}}$. This completes the proof of the relation (2.3).

Then, we prove the other direction $\overline{\mathcal{D}} \subset \mathcal{D}'$. By Theorem 2.2, we have

$$\mathcal{D} \subset \mathcal{D}'.$$

Hence, it remains to prove that the set $\mathcal{D}'$ is closed. Equivalently, we prove that $(\mathcal{D}')^c$ is open, where $(\mathcal{D}')^c$ is the complementary set with respect to $\mathbb{R}^{n \times n} \times \mathbb{R}^n$. Suppose that $(C, u^*) \in (\mathcal{D}')^c$. If $\|C\|_1 \neq 1$ or $\|u^*\|_1 \neq 1$, changing C and u^* by a small perturbation will not make $\|C\|_1 = \|u^*\|_1 = 1$. Now, we only consider the case when $\|C\|_1 = \|u^*\|_1 = 1$. Since $(C, u^*) \in (\mathcal{D}')^c$, the subgraph $\mathbb{G}_1(C, u^*)$ is connected and non-bipartite and the set $\mathcal{I}_{00}(C, u^*) = \emptyset$. Denote

$$\epsilon := \min \left\{ \min_{C_{ij} > 0} C_{ij}, \min_{u_i^* \neq 0} |u_i^*| \right\} > 0.$$

Suppose that we add a sufficiently small perturbation to the point (C, u^*) such that each component of C and u^* is changed by at most $\epsilon/2$. Then, all nonzero components of C and u^* are still nonzero after the perturbation. Therefore, the edges of the subgraph $\mathbb{G}_1(C, M^*)$ are not deleted after the perturbation and, thus, the subgraph is still connected and non-bipartite. Similarly, after perturbation, each node in $\mathcal{I}_0(C, M^*)$ either becomes nonzero or is connected to $\mathbb{G}_1(C, M^*)$, which implies that $\mathcal{I}_{00}(C, M^*)$ is still an empty set. Therefore, the perturbed instance still belongs to $(\mathcal{D}')^c$. Hence, the set $(\mathcal{D}')^c$ is open and we obtain the relation $\overline{\mathcal{D}} \subset \mathcal{D}'$. □

Using the results in Theorems 2.2 and 2.3, we provide an estimate on the scale of the new metric. Since $\mathcal{D}$ is a bounded set, there exists an upper bound on the minimum possible value of the complexity metric, which is defined below:

$$\mathbb{D}_\alpha^{min} := \min_{C \in \mathbb{S}_{+,1}^{n^2-1}, u^* \in \mathbb{S}_1^{n-1}} \mathbb{D}_\alpha(C, u^*).$$

The next theorem provides the expression of $\mathbb{D}_\alpha^{min}$.

Theorem 2.4 *Suppose that $n \geq 5$. Then, it holds that*

$$\mathbb{D}_\alpha^{min} = \begin{cases} \frac{n}{4\alpha} & \text{if } \alpha \leq \frac{n^2-3n-2}{n^2-5n+4} \\ \frac{n^2}{2(1-\alpha)(n-2)n+4\alpha} & \text{if } \frac{n}{n+2} \leq \alpha \leq \frac{n}{n+1} \\ \frac{n(n+1)}{2(1-\alpha)(n-2)(n+1)+4} & \text{if } \alpha \geq \frac{n}{n+1}. \end{cases}$$

In the regime $(n^2 - 3n - 2)/(n^2 - 5n + 4) \leq \alpha \leq n/(n+2)$, we have the estimate

$$\mathbb{D}_\alpha^{min} \in \left[\frac{n}{4\alpha}, \frac{n^2}{4\alpha(n-1)} \right].$$

The proof of Theorem 2.4 relies on the following two lemmas, which transform the computation of $\mathbb{D}_\alpha^{min}$ to a one-dimensional optimization problem. The first lemma upper-bounds the maximum possible distance.

Lemma 2.1 *Suppose that $n \geq 2$. It holds that*

$$\left(\mathbb{D}_\alpha^{min}\right)^{-1} \leq \max_{c \in \left[0, \frac{1}{n(n-1)}\right]} g(\alpha, c),$$

where the function $g(\alpha, c)$ is defined by

$$
\begin{aligned}
g(\alpha, c) := \min \Bigg\{ & 2(1 - \alpha) \cdot \frac{n-2}{n} + 4\alpha c, \quad 4\alpha(n-1)c, \\
& 2(1 - \alpha) \cdot \frac{n-4}{n} + 2\alpha \left(\frac{4}{n} - 4(n-2)c\right), \\
& 2(1 - \alpha) \cdot \frac{n-3}{n} + 2\alpha \left(\frac{3}{n} - (3n-5)c\right), \\
& 2(1 - \alpha) \cdot \frac{n-2}{n} + 2\alpha \left(\frac{2}{n} - 2(n-1)c\right), \\
& 2(1 - \alpha) \cdot \frac{n-1}{n} + 2\alpha \left(\frac{1}{n} - (n-1)c\right) \Bigg\}.
\end{aligned}
$$

We denote $g_i(\alpha, c)$ be the i-th term in the above minimization for all $i \in \{1, \ldots, 6\}$. The next lemma proves the other direction.

Lemma 2.2 *Suppose that $n \geq 2$. It holds that*

$$\left(\mathbb{D}_\alpha^{min}\right)^{-1} \geq \max_{c \in \left[0, \frac{1}{n(n-1)}\right]} g(\alpha, c),$$

where the function $g(\alpha, c)$ is defined in Lemma 2.1.

The proofs of both lemmas can be found in the appendix. Now, we provide the proof of Theorem 2.4.

Proof (Proof of Theorem 2.4) By the results of Lemmas 2.1 and 2.2, we only need to compute $\max_{c \in \left[0, \frac{1}{n(n-1)}\right]} g(\alpha, c)$. Let $\kappa := (1 - \alpha)/\alpha \in [0, +\infty]$. We study three cases below.

Case I. We first consider the case when $\kappa \geq 2(n-3)/[(n-4)(n-1)]$. We prove that $g(\alpha, c) = g_2(\alpha, c)$. Since $g_2(\alpha, c)$ has a larger gradient than $g_1(\alpha, c)$ and the function $g_i(\alpha, c)$ is decreasing in c for $i = 3, 4, 5, 6$, we only need to show that

$$(2.4) \qquad g_i\left(\alpha, \frac{1}{n(n-1)}\right) \geq g_2\left(\alpha, \frac{1}{n(n-1)}\right), \quad \forall i \in \{1, 3, 4, 5, 6\}.$$

The above inequality with $i = 1$ is equivalent to $\kappa \geq 2/(n-1)$, which is guaranteed by the assumption that $\kappa \geq 2(n-3)/[(n-4)(n-1)]$. For $i \in \{3, 4, 5, 6\}$, the inequality (2.4) is equivalent to

$$\kappa \geq \max \left\{ \frac{2(n-3)}{(n-1)(n-4)}, \frac{2(n-2)}{(n-1)(n-3)}, \frac{2}{n-2}, \frac{2}{n-1} \right\} = \frac{2(n-3)}{(n-1)(n-4)}.$$

Therefore, it holds that

$$g(\alpha, c) = g_2(\alpha, c) = 4\alpha(n-1)c.$$

whose maximum is attained at $c = [n(n-1)]^{-1}$ and

$$\max_{C,u^*} \mathbb{T}_\alpha(C, u^*) = g_2\left(\alpha, \frac{1}{n(n-1)}\right) = \frac{4\alpha}{n}.$$

Case II. Then, we consider the case when $\kappa \leq 2/n$. In this case, we prove that the maximum is achieved by the intersection point between $g_1(\alpha, c)$ (an increasing function in c) and $\min\{g_5(\alpha, c), g_6(\alpha, c)\}$ (a decreasing function in c). The intersection points between $g_1(\alpha, C)$ and the other five functions are

$$\frac{\kappa}{2n}, \quad \frac{2-\kappa}{n(2n-3)}, \quad \frac{3-\kappa}{3n(n-1)}, \quad \frac{1}{n^2}, \quad \frac{1+\kappa}{n(n+1)}.$$

In the regime $\kappa \leq 1/n$, we have

$$\frac{\kappa}{2n} \leq \frac{1+\kappa}{n(n+1)} \leq \frac{1}{n^2} \leq \min\left\{\frac{2-\kappa}{n(2n-3)}, \frac{3-\kappa}{3n(n-1)}\right\},$$

which implies that the maximum is attained at $c = (1+\kappa)/[n(n+1)]$. Hence, the maximum distance is

$$\max_{C,u^*} \mathbb{T}_\alpha(C, u^*) = g_1\left(\alpha, \frac{1+\kappa}{n(n+1)}\right) = \frac{2(1-\alpha)(n-2)(n+1)+4}{n(n+1)}.$$

In the regime $1/n \leq \kappa \leq 2/n$, we have

$$\frac{\kappa}{2n} \leq \frac{1}{n^2} \leq \frac{1+\kappa}{n(n+1)} \leq \min\left\{\frac{2-\kappa}{n(2n-3)}, \frac{3-\kappa}{3n(n-1)}\right\},$$

which implies that the maximum is attained at $c = 1/n^2$. Hence, the maximum distance is

$$\max_{C,u^*} \mathbb{T}_\alpha(C, u^*) = g_1\left(\alpha, \frac{1}{n^2}\right) = \frac{2(1-\alpha)(n-2)n+4\alpha}{n^2}.$$

Case III. We finally consider the case when $2/n \leq \kappa \leq 2(n-3)/[(n-4)(n-1)]$. In this regime, the intersection point between $g_2(\alpha, c)$ and $g_5(\alpha, c)$ is

$$\frac{\kappa(n-2)+2}{4n(n-1)} \leq \frac{\kappa}{2n}.$$

This implies that $g_2(\alpha, c)$ intersects with $g_5(\alpha, c)$ before $g_1(\alpha, c)$. Therefore, the maximum is attained at one of the intersects between $g_2(\alpha, c)$ and $g_i(\alpha, c)$ for $i = 3, 4, 5, 6$. By calculating the four intersects, the optimal c that achieves the maximum is given by

$$c^*(\kappa) := \min\left\{\frac{\kappa(n-4)+4}{n(6n-10)}, \frac{\kappa(n-3)+3}{n(5n-7)}, \frac{\kappa(n-2)+2}{4n(n-1)}, \frac{\kappa(n-1)+1}{3n(n-1)}\right\},$$

which is an increasing function in κ. If $\kappa = 2/n$, we can estimate that

(2.5)

$$
\begin{aligned}
&c^*(\kappa) \\
&= \min\left\{ \frac{2(n-4)/n+4}{n(6n-10)}, \frac{2(n-3)/n+3}{n(5n-7)}, \frac{2(n-2)/n+2}{4n(n-1)}, \frac{2(n-1)/n+1}{3n(n-1)} \right\} \\
&= \min\left\{ \frac{3n-4}{n^2(3n-5)}, \frac{5n-6}{n^2(5n-7)}, \frac{1}{n^2}, \frac{3n-2}{n^2(3n-3)} \right\} = \frac{1}{n^2}.
\end{aligned}
$$

Similarly, if $\kappa = 2(n-3)/[(n-4)(n-1)]$, it holds that

(2.6)
$$
c^*(\kappa) = \frac{1}{n(n-1)}.
$$

Combining (2.5) and (2.6), we have

$$
c^*(\kappa) \in \left[\frac{1}{n^2}, \frac{1}{n(n-1)} \right], \quad \forall \kappa \in \left[\frac{2}{n}, \frac{2(n-3)}{(n-4)(n-1)} \right].
$$

Therefore, the maximum distance satisfies the bound

$$
\max_{C,u^*} \mathbb{T}_\alpha(C, u^*) = g_2\left[\alpha, c^*(\kappa)\right] \in \left[\frac{4\alpha(n-1)}{n^2}, \frac{4\alpha}{n} \right].
$$

This completes the proof.   □

The results of Theorem 2.4 imply that in the regime where $\alpha \geq \Theta(1)$ and $1 - \alpha \geq \Theta(n^{-1})$, we have $\mathbb{D}_\alpha^{min} = O(n)$. This suggests that $n^{-1}\mathbb{D}_\alpha(C, u^*)$ may be a dimension-free complexity metric; see more examples supporting this conjecture in Section 3. In addition, the minimum possible value of the complexity is attained at

$$
\alpha^* := (n^2 - 5n + 4)/(n^2 - 3n - 2).
$$

Hence, the set of possible values of the complexity metric attains the maximum size by choosing $\alpha = \alpha^*$. This observation hints that α^* may be the optimal choice of α since it may enable the metric to differentiate instances with different complexities to the maximum degree. Using the exact formulation of $g(\alpha, c)$ in Lemma 2.1, we plot the minimum possible value of the complexity metric both without scaling and after scaling by n^{-1} in Figure 1. From the numerical results, we can see that the complexity scales with n if α is smaller than α^*, which is consistent with Theorem 2.4. If α is larger than α^*, the complexity metric for different values of n approximately lies on the same curve.

In the following theorem, we show that if $\alpha = \alpha^*$, the instances that attain the minimum value of the complexity metric are unique up to sign flips to components of the global solution.
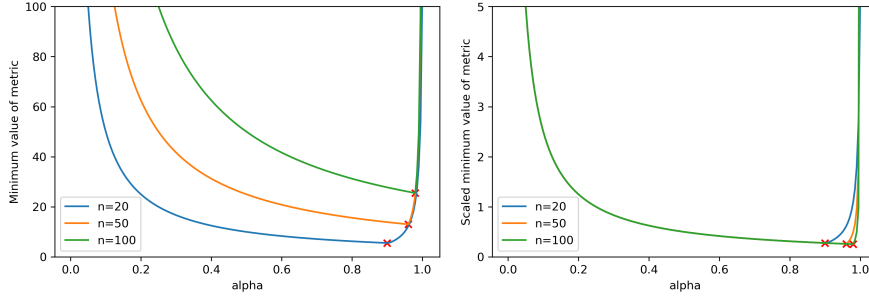
Fig. 1: Comparison of $\mathbb{D}_\alpha^{min}$ for $n = 20, 50, 100$. The red "×" sign refers to the value at α^*. In the right plot, the complexity metric is scaled by n^{-1}.

Theorem 2.5 *Suppose that $n \geq 5$ and the instance $\mathcal{MC}(C, u^*)$ satisfies*

$$\mathbb{D}_{\alpha^*}(C, u^*) = \frac{n}{4\alpha^*}.$$

Then, it holds that

$$|u_i^*| = \frac{1}{n}, \quad C_{ii} = 0, \quad \forall i \in [n]; \quad C_{ij} = \frac{1}{n(n-1)}, \quad \forall i, j \in [n], \ i \neq j.$$

Proof By the assumption that the complexity metric of (C, u^*) is finite, we have that $(C, u^*) \notin \overline{\mathcal{D}}$. It follows from Theorem 2.3 that the subset $\mathcal{I}_{00}(C, u^*)$ is empty and that $\mathbb{G}_1(C, u*)$ is connected and non-bipartite. Let $k := |\mathcal{I}_1(C, u^*)|$. For each node $i_0 \in \mathcal{I}_1(C, u^*)$, we define the new weight matrix $\tilde{C}$ as

$$\tilde{C}_{i_0 j} = \tilde{C}_{j i_0} = 0, \quad \forall j \in \mathcal{I}_1(C, u^*) \backslash \{i_0\};$$

$$\tilde{C}_{ij} = C_{ij} + \frac{2}{n^2 - 2(k-1)} \sum\nolimits_{j \in \mathcal{I}_1(C, u^*) \backslash \{i_0\}} C_{i_0 j}, \quad \text{otherwise.}$$

The subgraph $\mathbb{G}_1(\tilde{C}, u^*)$ is disconnected and, therefore, we have $(\tilde{C}, u^*) \in \overline{\mathcal{D}}$. It follows that

$$(2.7) \qquad \frac{4\alpha^*}{n} = [\mathbb{D}_{\alpha^*}(C, u^*)]^{-1} \leq \alpha^* \|C - \tilde{C}\|_1 = 4\alpha^* \sum\nolimits_{j \in \mathcal{I}_1(C, u^*) \backslash \{i_0\}} C_{i_0 j}.$$

For each node $i_0 \in \mathcal{I}_0(C, u^*)$, a similar construct of $\tilde{C}$ leads to

$$(2.8) \qquad \frac{4\alpha^*}{n} = [\mathbb{D}_{\alpha^*}(C, u^*)]^{-1} \leq 4\alpha^* \sum_{j \in \mathcal{I}_1(C, u^*)} C_{i_0 j}.$$

By summing inequality (2.7) over i_0 for all nodes in $\mathcal{I}_1(C, u^*)$ and summing inequality (2.8) over i_0 for all nodes in $\mathcal{I}_0(C, u^*)$, it follows that

$$(2.9) \qquad 4\alpha^* \leq 4\alpha^* \left[\sum_{i,j \in \mathcal{I}_1(C, u^*), i \neq j} C_{ij} + \sum_{i \in \mathcal{I}_1(C, u^*), j \in \mathcal{I}_0(C, u^*)} C_{ij} \right]$$

$$\leq 4\alpha^* \sum_{i,j \in [n], i \neq j} C_{ij} \leq 4\alpha^*,$$

where all inequalities should hold with equality. Since the last inequality in (2.9) holds with equality, we obtain that

$$C_{ii} = 0, \quad \forall i \in [n].$$

It follows from the equality of inequalities (2.7) and (2.8) that

(2.10)
$$\sum_{j \in \mathcal{I}_1(C, u^*) \setminus \{i\}} C_{ij} = \frac{1}{n}, \ \forall i \in \mathcal{I}_1(C, u^*); \qquad \sum_{j \in \mathcal{I}_1(C, u^*)} C_{ij} = \frac{1}{n}, \ \forall i \in \mathcal{I}_0(C, u^*).$$

Using the condition that $\|C\|_1 = 1$, the above equalities imply that all weights of C are limited to edges with a node in $\mathcal{I}_1(C, u^*)$. Namely, we have

(2.11)
$$\sum_{j \in \mathcal{I}_0(C, u^*)} C_{ij} = 0, \quad \forall i \in \mathcal{I}_1(C, u^*).$$

If $\mathcal{I}_o(C, u^*)$ is not empty, the above equality contradicts the second equality in (2.10). Hence, the point (C, u^*) satisfies that $\mathcal{I}_0(C, u^*) = \emptyset$. By a similar analysis of the bipartite instance in Lemma 2.1, for every 4-element subset $\{i, j, k, \ell\}$ of $[n]$, it holds that

$$2(1 - \alpha^*)(1 - |u_i^*| - |u_j^*| - |u_k^*| - |u_\ell^*|) + 4\alpha^*(C_{ij} + C_{k\ell}) = 4\alpha^*/n.$$

Taking the average of the above equality over $\{i, j, k, \ell\}$ for all 4-element subsets of $[n-1]$, we obtain that

$$2(1 - \alpha^*)\left(1 - \frac{3\|u_{1:n-1}^*\|_1}{n - 1}\right) + 4\alpha^* \frac{2}{(n-1)(n-2)}\|C_{1:n-1, 1:n-1}\|_1 = \frac{4\alpha^*}{n}.$$

Using the first equality in (2.10) and the symmetry of C, it holds that $\|C_{1:n-1, 1:n-1}\|_1 = 1 - 2/n$. Substituting into the above equality, we know

$$2(1 - \alpha^*)\left(1 - \frac{3\|u_{1:n-1}^*\|_1}{n - 1}\right) = 4\alpha^* \cdot \frac{n - 3}{n(n-1)}.$$

By recalling that $\alpha^* = (n - 1)(n - 4)/(n^2 - 3n - 2)$, the above inequality leads to

$$\|u_{1:n-1}^*\|_1 = (n - 1)/n,$$

which is equivalent to $|u_n^*| = 1/n$. By the same proof technique, we conclude that

$$|u_i^*| = 1/n, \quad \forall i \in [n].$$

By substituting back into equality (2.11), it holds for all 4-element subsets $\{i, j, k, \ell\} \subset [n]$ that

$$C_{ij} + C_{k\ell} = \frac{2}{n(n-1)},$$

which implies that

$$C_{ij} = \frac{1}{n(n-1)}, \quad \forall i, j \in [n] \ \text{ s.t. } \ i \neq j.$$

$\square$

The above theorem states that if we choose the weight $\alpha = \alpha^*$, the "easiest" instance is unique up to a change in the signs of the components of the global solution u^*. In the next theorem, we show that a similar property as α^* holds if we set α to be

$$\alpha^\diamond := n/(n+2).$$

Theorem 2.6 *Suppose that $n \geq 5$ and the instance $\mathcal{MC}(C, u^*)$ satisfies*

$$\mathbb{D}_{\alpha^\diamond}(C, u^*) = \mathbb{D}_\alpha^{min} = \frac{n(n+2)}{4(n-1)}.$$

Then, it holds that

$$|u_i^*| = \frac{1}{n}, \quad \forall i \in [n]; \quad C = \frac{1}{n^2} I_n.$$

Since the proof is similar to that of Theorem 2.5, we omit it for brevity. The above theorem implies that the weight matrix C of the "easiest" instances is a constant multiple of the identity matrix I_n, which satisfies the δ-RIP$_{2,2}$ condition with $\delta = 0$. This is consistent with the common sense that the RIP constant δ being 0 is the optimal situation. Hence, Theorem 2.6 suggests that the choice $\alpha^\diamond = n/(n+2)$ may potentially be the optimal choice of α. On the other hand, we will prove in Section 4.1 that the "easiest" instances in Theorems 2.5 and 2.6 all have a benign landscape in the sense that they satisfy the strict-saddle property [45], which guarantees the polynomial-time global convergence of various algorithms. If the weight α is different from α^* and $\alpha^\diamond$, there may exist multiple "essentially" different instances attaining the minimum complexity.

3 Connections to existing results

In this section, we provide estimates of the proposed complexity metric on well-studied problem instances. More specifically, we consider matrix sensing problems satisfying the RIP condition and matrix completion problems under the Bernoulli model. In addition, we construct a class of instances parameterized by a single parameter. We estimate the threshold of the parameter that separates instances with a desirable optimization landscape from those with a bad landscape.

3.1 Matrix sensing problem: RIP condition

We first consider instances of problem (1.7) that satisfy the δ-RIP$_{2,2}$ condition, where $\delta \in [0, 1)$ is the RIP constant. However, the constraint that $C \in \mathbb{S}_{+,1}^{n^2-1}$ is inconsistent with the RIP condition (1.3) in the sense that the entries of C are averagely on the scale of n^{-2}, but the RIP condition requires that the entries of C be on the scale of $O(1)$. Therefore, we generalize the definition of the RIP condition to deal with the inconsistent scaling:

Definition 3.1 Given natural numbers r and s, the function $f(\cdot; M^*)$ is said to satisfy the **Restricted Isometry Property** (RIP) of rank $(2r, 2s)$ for a constant $\delta \in [0, 1)$, denoted as δ-RIP$_{2r,2s}$, if there exist constants $c_1, c_2 \geq 0$ such that $c_2/c_1 = (1 + \delta)/(1 - \delta)$ and

$$(3.1) \qquad c_1 \|K\|_F^2 \leq \left[\nabla^2 f(M; M^*) \right] (K, K) \leq c_2 \|K\|_F^2$$

holds for all matrices $M, K \in \mathbb{R}^{n \times n}$ such that $\text{rank}(M) \leq 2r, \text{rank}(K) \leq 2s$.

The above definition of the RIP condition is scale-free in the sense that for any constant $c > 0$, the function $cf(\cdot; M^*)$ satisfies the δ-RIP$_{2r,2s}$ condition if and only if $f(\cdot; M^*)$ satisfies the same condition. The following lemma transforms the RIP$_{2,2}$ condition to a condition on the weight matrix C.

Lemma 3.1 *Given a constant $\delta \in [0, 1)$, the instance $\mathcal{MC}(C, u^*)$ satisfies the δ-RIP$_{2,2}$ condition if and only if*

$$\frac{\min_{i,j \in [n]} C_{ij}}{\max_{i,j \in [n]} C_{ij}} \geq \frac{1 - \delta}{1 + \delta}.$$

The proof of Lemma 3.1 can be found in the appendix. Since the instances satisfying the RIP condition have a benign optimization landscape, we expect that the complexity metric is upper-bounded for those instances. With the help of Lemma 3.1, we provide an upper bound on the complexity metric of every instance satisfying the δ-RIP$_{2,2}$ condition.

Theorem 3.1 *Suppose that $\delta \in [0, 1)$ is a constant and the instance $\mathcal{MC}(C, u^*)$ satisfies the δ-RIP$_{2,2}$ condition. Then, it holds that*

$$\mathbb{D}_\alpha(C, u^*) \leq \frac{n^2(1 + \delta) - 2\delta}{2\alpha(1 - \delta)}.$$

The minimum distance is attained by the instance $\mathcal{MC}(C^\delta, u^\delta)$, where

$$C_{11}^\delta = \frac{1 - \delta}{(1 + \delta)n^2 - 2\delta}; \quad C_{ij}^\delta = \frac{1 + \delta}{(1 + \delta)n^2 - 2\delta}, \quad \forall i, j \in [n] \times [n] \backslash \{(1, 1)\};$$

$$u_1^\delta = 1; \quad u_i^\delta = 0, \quad \forall i \geq 2.$$

Proof Suppose that $(\tilde{C}, \tilde{u}^*) \in \overline{\mathcal{D}}$ is the closest pair in $\overline{\mathcal{D}}$ to (C, u^*). By Theorem 2.3, at least one entry of $\tilde{C}$ is 0. Suppose that $\tilde{C}_{i_0 j_0} = 0$ for some indices $i_0, j_0 \in [n]$. Then, the distance between C and $\tilde{C}$ satisfies

$$\|C - \tilde{C}\|_1 \geq 2C_{i_0 j_0}.$$

By Lemma 3.1, the minimum value of $C_{i_0 j_0}$ can be estimated by

$$1 = \sum_{i,j \in [n]} C_{ij} \leq C_{i_0 j_0} + \sum_{(i,j) \neq (i_0, j_0)} \frac{1 + \delta}{1 - \delta} \cdot C_{i_0 j_0} = \frac{(n^2 - 1)(1 + \delta) + 1}{1 - \delta} \cdot C_{i_0 j_0}.$$

It follows that

$$C_{i_0 j_0} \geq \frac{1 - \delta}{n^2(1 + \delta) - 2\delta}.$$

Therefore, the distance between (C, u^*) and $(\tilde{C}, \tilde{u}^*)$ satisfies

$$\alpha \|C - \tilde{C}\|_1 + (1 - \alpha)\|u^* - \tilde{u}^*\|_1 \geq 2\alpha C_{i_0 j_0} \geq \frac{2\alpha(1 - \delta)}{n^2(1 + \delta) - 2\delta}.$$

The upper bound on $\mathbb{D}_\alpha(C, u^*)$ stated in the theorem can be derived by taking the inverse of the above inequality. For the instance $\mathcal{MC}(C^\delta, u^\delta)$, the closest parameters in $\overline{\mathcal{D}}$ can be constructed by setting C_{11}^δ to 0 and normalizing the remaining entries of C.   □

We note that the upper bound on $\mathbb{D}_\alpha(C, u^*)$ is increasing in δ, which is consistent with the intuition that a smaller δ will lead to a better optimization landscape.

By suitably generalizing the definitions of $\mathbb{D}_\alpha(C, u^*)$ and $\mathcal{D}$, the results in Theorem 3.1 can be extended to the general problem (1.2) in the rank-1 case with a slightly weaker bound. In this work, we provide one possible generalization. For the problem (1.2), each instance is defined by the loss function $f(\cdot; \cdot)$ and the ground truth M^*, which we denote as (f, M^*). We assume that the loss function satisfies

(3.2) $f(M^*; M^*) = \min_{K \in \mathbb{R}^{n \times n}} f(K; M^*),$

$$\forall M^* \in \mathbb{R}^{n \times n} \quad \text{s.t.} \quad M^* \succeq 0, \text{rank}(M^*) = 1.$$

In the special case when $f(\cdot; \cdot)$ is the weighted ℓ_2-loss function in (1.7), the above condition implies that $C_{ij} \geq 0$ for all $i, j \in [n]$. Similar to the normalization constraint $C \in \mathbb{S}_{+,1}^{n^2-1}$, we consider any objective function $f(\cdot; M^*)$ with the property that

(3.3) $$\sum_{i,j \in [n]} [f(M^* + E_{ij}; M^*) - f(M^*; M^*)] = 1,$$

where E_{ij} is defined in (C.1). For the normalization constraint $u^* \in \mathbb{S}_1^{n-1}$, we consider any global truth with the property that

(3.4) $$\sum_{i \in [n]} \sqrt{M_{ii}^*} = 1.$$

The set of degenerate instances is given by

$$\mathcal{D} := \left\{ (f, M^*) \;\middle|\; f(\cdot; \cdot) \text{ and } M^* \text{ satisfy } (3.2)\text{-}(3.4), \right.$$

$$\left. \exists M \neq M^* \quad \text{s.t.} f(M; M^*) = f(M^*; M^*), M^* \succeq 0, \text{rank}(M^*) = 1 \right\}.$$

The "entry-wise ℓ_1-norm" between two arbitrary functions $h^1(\cdot)$ and $h^2(\cdot)$ with the domain $\mathbb{R}^{n \times n}$ is defined as the restricted ℓ_∞-Lipschitz constant of $h^1 - h^2$. Namely, we define $\|h^1 - h^2\|_1$ to be

$$\|h^1 - h^2\|_1 := \sup_{K, L \in \mathbb{R}^{n \times n}} \frac{\left| \left(h^1(K) - h^2(K) \right) - \left(h^1(L) - h^2(L) \right) \right|}{\max_{i, j \in [n]} (K_{ij} - L_{ij})^2}$$

$$\text{s.t.} \quad K \neq L, \quad \mathrm{rank}(K - L) \leq 2.$$

For every constant $\alpha \in [0, 1]$, the distance between two instances (f, M^*) and $(\tilde{f}, \tilde{M}^*)$ is defined as

$$\mathrm{dist}_\alpha \left[(f, M^*), (\tilde{f}, \tilde{M}^*) \right] := \alpha \|f(\cdot; M^*) - \tilde{f}(\cdot; \tilde{M}^*)\|_1 + (1 - \alpha)\|u^* - \tilde{u}^*\|_1,$$

where $u^*, \tilde{u}^* \in \mathbb{R}^n$ satisfy $u^*(u^*)^T = M^*$ and $\tilde{u}^*(\tilde{u}^*)^T = \tilde{M}^*$. Finally, the complexity metric is given by

$$\mathbb{D}_\alpha(f, M^*) := \left[\inf_{(\tilde{f}, \tilde{M}^*) \in \mathcal{D}} \mathrm{dist}_\alpha \left[(f, M^*), (\tilde{f}, \tilde{M}^*) \right] \right]^{-1}.$$

We note that the definitions of $\mathcal{D}$ and $\mathbb{D}_\alpha(f, M^*)$ are consistent with those of instance (1.7). The following theorem provides an upper bound on the complexity metric of any instance satisfying the $\mathrm{RIP}_{2,2}$ condition.

Theorem 3.2 *Let $\alpha \in [0, 1]$ and $\delta \in [0, 1)$ be two constants. Suppose that the function $f(\cdot; M^*)$ satisfies the δ-$\mathrm{RIP}_{2,2}$ condition and the normalization constraint (3.3). Then, it holds that*

$$\mathbb{D}_\alpha(f, M^*) \leq \frac{n^2(1 + \delta)}{\alpha(1 - \delta)}$$

Proof We fix the instance (f, M^*) and assume that $(\tilde{f}, \tilde{M}^*) \in \mathcal{D}$. Suppose that the matrix $M \neq \tilde{M}^*$ satisfies

$$\sum_{i \in [n]} \sqrt{M_{ii}} = 1, \quad \tilde{f}(M; \tilde{M}^*) = \tilde{f}(\tilde{M}^*; \tilde{M}^*).$$

We first consider the case when $M \neq M^*$. In this case, we can estimate that

(3.5) $\quad \|f(\cdot; M^*) - \tilde{f}(\cdot; \tilde{M}^*)\|_1$

$$\geq \frac{\left| \left[f(M; M^*) - \tilde{f}(M; \tilde{M}^*) \right] - \left[f(M^*; M^*) - \tilde{f}(M^*; \tilde{M}^*) \right] \right|}{\max_{i, j \in [n]} (M_{ij} - M_{ij}^*)^2}$$

$$= \frac{\left| [f(M; M^*) - f(M^*; M^*)] + \left[\tilde{f}(M^*; \tilde{M}^*) - \tilde{f}(M; \tilde{M}^*) \right] \right|}{\max_{i, j \in [n]} (M_{ij} - M_{ij}^*)^2}$$

$$= \frac{\left| [f(M; M^*) - f(M^*; M^*)] + \left[\tilde{f}(M^*; \tilde{M}^*) - \tilde{f}(\tilde{M}^*; \tilde{M}^*) \right] \right|}{\max_{i, j \in [n]} (M_{ij} - M_{ij}^*)^2}$$

$$\geq \frac{f(M; M^*) - f(M^*; M^*)}{\max_{i, j \in [n]} (M_{ij} - M_{ij}^*)^2} \geq \frac{(c_1/2) \cdot \|M - M^*\|_F^2}{\max_{i, j \in [n]} (M_{ij} - M_{ij}^*)^2} \geq \frac{c_1}{2},$$

where c_1 is the constant in the RIP condition of $f(\cdot; M^*)$. The second inequality is due to

$$f(M; M^*) - f(M^*; M^*) \geq 0, \quad \tilde{f}(M^*; \tilde{M}^*) - \tilde{f}(\tilde{M}^*; \tilde{M}^*) \geq 0.$$

The second last inequality follows from the global optimality of M^* and the second inequality after inequality (12) in [57], namely,

$$f(M; M^*) \geq f(M^*; M^*) + \frac{c_1}{2}\|M - M^*\|_F^2, \quad \forall M \in \mathbb{R}^{n \times n}, \text{ rank}(M) \leq 1.$$

Now, we provide a lower bound on c_1. Using the normalization constraint (3.3) and the stationarity of M^*, it holds that

$$1 = \sum_{i,j \in [n]} [f(M^* + E_{ij}; M^*) - f(M^*; M^*)] \leq \frac{c_2}{2} \cdot \sum_{i,j \in [n]} \|E_{ij}\|_F^2 = \frac{c_2 n^2}{2},$$

which implies that $c_2 \geq 2n^{-2}$. Using the relation $c_2/c_1 = (1+\delta)/(1-\delta)$, we obtain that

$$c_1 \geq \frac{2(1-\delta)}{n^2(1+\delta)}.$$

By substituting into inequality (3.5), it follows that

$$\|f(\cdot; M^*) - \tilde{f}(\cdot; \tilde{M}^*)\|_1 \geq \frac{1-\delta}{n^2(1+\delta)}.$$

which leads to $\text{dist}_\alpha[(f, M^*), (\tilde{f}, \tilde{M}^*)] \geq \alpha(1-\delta)/[n^2(1+\delta)]$. Now, the desired bound on $\mathbb{D}_\alpha(f, M^*)$ follows from taking the inverse. In the case when $M = M^*$, we can replace M with $\tilde{M}^*$ and the proof can be done in the same way. $\quad\square$

We note that in Theorems 3.1 and 3.2, the normalization constraint on u^* is not used. Therefore, the upper bound may be loose for certain instances since the information about the global solution is not used. On the other hand, in the case when $\alpha(1-\delta) = \Theta(1)$, the upper bound is on the order of $O(n^2)$, which is $O(n)$ larger than the minimum possible complexity metric in Theorem 2.4. Now, we provide a remedy to the aforementioned issue for problem (1.7). With the knowledge about the incoherence of the global solution, we can improve the upper bound on the complexity metric.

Theorem 3.3 *Suppose that the instance $\mathcal{MC}(C, u^*)$ satisfies the δ-RIP$_{2,2}$ condition and u^* satisfies the μ-incoherence condition. Then, it holds that*

$$\mathbb{D}_\alpha(C, u^*) \leq \max\left\{\frac{n(1+\delta)}{4\alpha(1-\delta)}, \frac{1}{2(1-\alpha)\mu}\right\} \times \min\left\{\left(\frac{1}{\mu} - \frac{1}{n}\right)^{-1}, 3\mu\right\}.$$

Before proving the estimation of the complexity metric, we prove two properties of μ-incoherent vectors.

Lemma 3.2 *Given any constant $\mu \in [1, n]$, suppose that u^* satisfies the μ-incoherence condition and $\|u^*\|_1 = 1$. Then, the following properties hold:*

1. u^* has at least n/μ nonzero components;
2. $|u_i^*| \leq \mu/n$ for all $i \in [n]$.

The proof of Lemma 3.2 can be found in the appendix. We first lower-bound the perturbation of the weight matrix C.

Lemma 3.3 *Suppose that the instance $\mathcal{MC}(C, u^*)$ satisfies the δ-RIP$_{2,2}$ condition and the weight matrix $\tilde{C} \in \mathbb{S}_{+,1}^{n^2-1}$ has N zero entries, where $\delta \in [0, 1)$ and $N \in [n^2]$. Then, it holds that*

$$\|C - \tilde{C}\|_1 \geq 2\sum\nolimits_{(i,j)\in\mathcal{N}} C_{ij} \geq \frac{2(1-\delta)N}{(1+\delta)n^2 - 2\delta N},$$

where $\mathcal{N}$ is the set of indices of zero entries of $\tilde{C}$.

Proof The δ-RIP$_{2,2}$ condition implies that

$$\frac{\min_{i,j} C_{ij}}{\max_{i,j} C_{ij}} \geq \frac{1-\delta}{1+\delta}.$$

Therefore, considering the average of entries in $\mathcal{N}$ and that of entries not in $\mathcal{N}$, we have

$$\frac{\frac{1}{N}\sum_{(i,j)\in\mathcal{N}} C_{ij}}{\frac{1}{n^2-N}\sum_{(i,j)\notin\mathcal{N}} C_{ij}} \geq \frac{1-\delta}{1+\delta},$$

which further leads to

$$\sum_{(i,j)\in\mathcal{N}} C_{ij} \geq \frac{1-\delta}{1+\delta} \cdot \frac{N}{n^2-N}\sum_{(i,j)\notin\mathcal{N}} C_{ij} = \frac{1-\delta}{1+\delta} \cdot \frac{N}{n^2-N}\left(1 - \sum_{(i,j)\in\mathcal{N}} C_{ij}\right).$$

The above inequality is equivalent to

$$\sum\nolimits_{(i,j)\in\mathcal{N}} C_{ij} \geq \frac{(1-\delta)N}{(1+\delta)n^2 - 2\delta N}.$$

Hence, the distance between C and $\tilde{C}$ is lower-bounded as

$$\|C - \tilde{C}\|_1 \geq 2\sum\nolimits_{(i,j)\in\mathcal{N}} C_{ij} \geq \frac{2(1-\delta)N}{(1+\delta)n^2 - 2\delta N}.$$

This completes the proof. $\square$

Proof (Proof of Theorem 3.3) Suppose that $\mathcal{MC}(\tilde{C}, \tilde{u}^*) \in \overline{\mathcal{D}}$ is the instance such that

$$[\mathbb{D}_\alpha(C, u^*)]^{-1} = \alpha\|C - \tilde{C}\|_1 + (1-\alpha)\|u^* - \tilde{u}^*\|_1.$$

In the following, we split the proof into two steps.

Step I. We first fix $\tilde{u}^*$ and consider the closest matrix $\tilde{C}$ to C such that $(\tilde{C}, \tilde{u}^*) \in \overline{\mathcal{D}}$. Let $k := |\mathcal{I}_1(\tilde{C}, \tilde{u}^*)|$. Without loss of generality, we assume that

$$\mathcal{I}_1(\tilde{C}, \tilde{u}^*) = \{1, \ldots, k\}, \quad \mathcal{I}_0(\tilde{C}, \tilde{u}^*) = \{k+1, \ldots, n\}.$$

We first consider the case when $k \geq 2$. If $\mathbb{G}_1(\tilde{C}, \tilde{u}^*)$ is disconnected, at least $2(k-1)$ entries of $\tilde{C}$ are 0. If $\mathbb{G}_1(\tilde{C}, \tilde{u}^*)$ are bipartite, at least $k^2/2 \geq 2(k-1)$ entries of $\tilde{C}$ are 0. If $\mathcal{I}_{00}(\tilde{C}, \tilde{u}^*)$ is non-empty, at least $2k$ entries of $\tilde{C}$ are 0. Otherwise if $k = 1$, at least one entry of $\tilde{C}$ should be 0 to make $\mathbb{G}_1(\tilde{C}, \tilde{u}^*)$ bipartite. In summary, at least $N(k)$ entries of $\tilde{C}$ are 0, where

$$N(k) := \max\{2(k-1), 1\}.$$

Using the results in Lemma 3.3, the distance between C and $\tilde{C}$ is at least

$$(3.6) \qquad \|C - \tilde{C}\|_1 \geq \frac{2(1-\delta)N(k)}{(1+\delta)n^2 - 2\delta N(k)}.$$

We note that the distance is monotonously increasing as a function of k.

Step II. Now, we consider the optimal choice of $\tilde{u}^*$ based on the lower bound in (3.6). Let

$$\ell := |\mathcal{I}_1(C, u^*)|, \quad k := |\mathcal{I}_1(\tilde{C}, \tilde{u}^*)|.$$

Since the distance between C and $\tilde{C}$ is a monotonously increasing function of k, the minimum distance between (C, u^*) and $(\tilde{C}, \tilde{u}^*)$ cannot be attained by $k > \ell$. Therefore, we focus on the case when $k \leq \ell$. Without loss of generality, we assume that

$$|u_1^*| \geq |u_2^*| \geq \cdots \geq |u_\ell^*| > 0; \quad |u_i^*| = 0, \quad \forall i \geq \ell + 1.$$

Then, the distance between u^* and $\tilde{u}^*$ satisfies

$$(3.7) \qquad \|u^* - \tilde{u}^*\|_1 \geq 2\sum\nolimits_{i=k+1}^{\ell} |u_i^*|.$$

Denote the distance between (C, u^*) and $(\tilde{C}, \tilde{u}^*)$ by

$$d_\alpha := \alpha\|C - \tilde{C}\|_1 + (1-\alpha)\|u^* - \tilde{u}^*\|_1.$$

Step II-1. We first consider the case when $\mu \leq 2n/3$. Combining inequalities (3.6) and (3.7), we obtain a lower bound on d_α:

$$d_\alpha \geq \min_{k \in [\ell]} \left[\frac{2\alpha(1-\delta)N(k)}{n^2(1+\delta) - 2\delta N(k)} + 2(1-\alpha)\sum\nolimits_{i=k+1}^{\ell} |u_i^*| \right].$$

For every $k \in [\ell]$, the term inside the above minimization can be lower-bounded by

$$\frac{2\alpha(1-\delta)N(k)}{n^2(1+\delta) - 2\delta N(k)} + 2(1-\alpha)\sum\nolimits_{i=k+1}^{\ell} |u_i^*|$$
$$\geq \frac{2\alpha(1-\delta) \cdot 2(k-1)}{n^2(1+\delta)} + 2(1-\alpha)\sum\nolimits_{i=k+1}^{\ell} |u_i^*|$$
$$= \frac{4\alpha(1-\delta)}{n^2(1+\delta)} \cdot (k-1) + 2(1-\alpha)\sum\nolimits_{i=k+1}^{\ell} |u_i^*|.$$

The minimum of the right-hand side over $k \in [\ell]$ can be solved in closed form and is equal to

$$\sum\nolimits_{i=2}^{\ell} \min \left\{ \frac{4\alpha(1-\delta)}{n^2(1+\delta)}, 2(1-\alpha)|u_i^*| \right\}.$$

Using the second property in Lemma 3.2, we have

$$\min \left\{ \frac{4\alpha(1-\delta)}{n^2(1+\delta)}, 2(1-\alpha)|u_i^*| \right\} \geq \min \left\{ \frac{4\alpha(1-\delta)}{n^2(1+\delta)} \cdot \frac{n|u_i^*|}{\mu}, 2(1-\alpha)|u_i^*| \right\}$$

$$= \min \left\{ \frac{4\alpha(1-\delta)}{\mu n(1+\delta)}, 2(1-\alpha) \right\} \cdot |u_i^*|.$$

Taking the summation over $k \in \{2, \ldots, \ell\}$, we can conclude that

$$(3.8) \qquad d_\alpha \geq \sum\nolimits_{k=2}^{\ell} \min \left\{ \frac{4\alpha(1-\delta)}{\mu n(1+\delta)}, 2(1-\alpha) \right\} \cdot |u_i^*|$$

$$= \min \left\{ \frac{4\alpha(1-\delta)}{\mu n(1+\delta)}, 2(1-\alpha) \right\} \cdot \sum\nolimits_{k=2}^{\ell} |u_i^*|.$$

Using the second property in Lemma 3.2 and $\|u^*\|_1 = 1$, it follows that

$$\sum\nolimits_{k=2}^{\ell} |u_i^*| \geq 1 - \frac{\mu}{n}.$$

Substituting back into inequality (3.8), we have

$$d_\alpha \geq \min \left\{ \frac{4\alpha(1-\delta)}{\mu n(1+\delta)}, 2(1-\alpha) \right\} \cdot \left(1 - \frac{\mu}{n} \right).$$

Step II-2. Next, we consider the case when $\mu \geq 2n/3$. By Theorem 3.1, the distance is at least

$$d_\alpha \geq \frac{2\alpha(1-\delta)}{n^2(1+\delta) - 2\delta} \geq \frac{2\alpha(1-\delta)}{(3/2)\mu \cdot n(1+\delta)} \geq \min \left\{ \frac{4\alpha(1-\delta)}{\mu n(1+\delta)}, 2(1-\alpha) \right\} \cdot \frac{1}{3},$$

where the second inequality is due to the assumption that $\mu \geq 2n/3$.

By combining *Steps II-1* and *II-2*, the distance is lower-bounded by

$$d_\alpha \geq \min \left\{ \frac{4\alpha(1-\delta)}{\mu n(1+\delta)}, 2(1-\alpha) \right\} \times \max \left\{ 1 - \frac{\mu}{n}, \frac{1}{3} \right\}$$

$$= \min \left\{ \frac{4\alpha(1-\delta)}{n(1+\delta)}, 2(1-\alpha)\mu \right\} \times \max \left\{ \frac{1}{\mu} - \frac{1}{n}, \frac{1}{3\mu} \right\}.$$

The proof is completed by using the relation between d_α and $\mathbb{T}_\alpha(C, u^*)$. $\qquad\square$

From the above theorem, we can use the weight α to control the balance between the RIP constant δ and the incoherence μ. If we choose $1 - \alpha = \Theta(n^{-1})$, then the complexity can be upper-bounded by

$$\mathbb{D}_\alpha(C, u^*) = \mu n \cdot \max\left\{O\left(\frac{1+\delta}{1-\delta}\right), O\left(\frac{1}{\mu}\right)\right\} = O\left(\mu n \cdot \frac{1+\delta}{1-\delta}\right).$$

In addition, if it holds that $\mu = O(1)$ and $(1-\delta)^{-1} = O(1)$, then the complexity is upper-bounded by $O(n)$, which matches the minimum possible complexity in Theorem 2.4 up to a constant. Although the complexity metric may have a large value for extreme instances (i.e., instances with a large incoherence), the complexity of regular instances achieves the optimal value up to a constant. We conclude the discussion of instances with the RIP condition by showing that the dependence of δ in Theorem 3.3 is tight up to a constant.

Theorem 3.4 *Suppose that $n \geq 4$, $\alpha \in [0, 1]$, $\mu \in [1, n]$ and $\delta \in [0, 1)$. Let $\ell := \lceil n/\mu \rceil$. Then, there exists an instance $\mathcal{MC}(C, u^*)$ such that $\mathcal{MC}(C, u^*)$ satisfies the δ-RIP$_{2,2}$ condition and*

$$\mathbb{D}_\alpha(C, u^*) \geq \frac{n(1+\delta)}{4\alpha(1-\delta)} \cdot \min\left\{\frac{n\mu}{\mu\ell - \mu}, \mu\right\}.$$

Proof The proof is split into two different cases.

Case I. We first consider the case when $\mu \leq n/2$. We construct the weight matrix $\tilde{C}$ as

$$\tilde{C}_{1i} = \tilde{C}_{i1} = 0, \quad \forall i \in \{2, \ldots, \ell\}; \quad \tilde{C}_{ij} = \frac{1}{n^2 - 2(\ell - 1)}, \quad \text{otherwise.}$$

For the instance $\mathcal{MC}(\tilde{C}, u^*)$, node 1 is disconnected from nodes $\{2, \ldots, \ell\}$ and thus, the subgraph $\mathbb{G}_1(\tilde{C}, u^*)$ is disconnected. This implies that $(\tilde{C}, u^*) \in \overline{\mathcal{D}}$. The matrix C is defined as

$$C_{1i} = C_{i1} = \frac{1 - \delta}{(1+\delta)n^2 - 4\delta(\ell - 1)}, \quad \forall i \in \{2, \ldots, \ell\};$$
$$C_{ij} = \frac{1 + \delta}{(1+\delta)n^2 - 4\delta(\ell - 1)}, \quad \text{otherwise.}$$

By Lemma 3.1, the weight matrix C ensures that $\mathcal{MC}(C, u^*)$ satisfies the δ-RIP$_{2,2}$ condition. The complexity of $\mathcal{MC}(C, u^*)$ is lower-bounded by

$$\mathbb{D}_\alpha(C, u^*) \geq \left(\alpha\|C - \tilde{C}\|_1\right)^{-1} = \frac{(1+\delta)n^2 - 4\delta(\ell - 1)}{4\alpha(\ell - 1)(1-\delta)}$$
$$\geq \frac{(1+\delta)(n^2 - 2n)}{4\alpha(\ell - 1)(1-\delta)} = \frac{n(1+\delta)}{4\alpha(1-\delta)} \cdot \frac{n-2}{\ell - 1} \geq \frac{n(1+\delta)}{4\alpha(1-\delta)} \cdot \frac{n\mu}{2(n\ell - 1)},$$

where the second last inequality follows from $4\delta \leq 2(1+\delta)$ and the last inequality is due to $n \geq 4$.

Case II. Next, we consider the case when $\mu \geq n/2$. Theorem 3.1 implies that there exists an instance $\mathcal{MC}(C, u^*)$ such that

$$\mathbb{D}_\alpha(C, u^*) = \frac{n^2(1 + \delta) - 2\delta}{2\alpha(1 - \delta)} \geq \frac{(n^2 - 1)(1 + \delta)}{2\alpha(1 - \delta)} \geq \frac{n(1 + \delta)}{2\alpha(1 - \delta)} \cdot \frac{n}{2},$$

where the first inequality results from $2\delta \leq 1 + \delta$ and the second inequality is in light of $n \geq 4$. Using the condition that $\mu \leq n$, it follows that

$$\mathbb{D}_\alpha(C, u^*) \geq \frac{n(1 + \delta)}{4\alpha(1 - \delta)} \cdot \mu.$$

Combining *Cases I* and *II* completes the proof. $\qquad\qquad\square$

3.2 Matrix completion problem: Bernoulli model and incoherence condition

Next, we consider instances $\mathcal{MC}(C, u^*)$ of problem (1.7) where the global solution u^* is μ-incoherent and the random weight matrix C obeys the Bernoulli model. Similar to the RIP condition, we need to generalize the definition of the Bernoulli model under the normalization constraint.

Definition 3.2 Given the sampling rate $p \in (0, 1]$, a random matrix $C \in \mathbb{S}_{+,1}^{n^2-1}$ is said to obey the **Bernoulli model** if

$$C_{ij} = \frac{\delta_{ij}}{\sum_{k,\ell \in [n]} \delta_{k\ell}}, \quad \forall i, j \in [n],$$

where $\{\delta_{k\ell} | k, \ell \in [n]\}$ are independent Bernoulli random variables with the parameter p.

We note that the above model is well defined only when $\sum_{i,j} \delta_{ij} > 0$, which happens with probability $1 - (1 - p)^{n^2} \geq 1 - \exp(-n^2 p)$. This probability is sufficiently large if $n^2 p \gg 1$. In [15], the authors showed that $p \geq \Theta(\mu \log n/n)$ is necessary and under this condition, the success probability is at least $1 - O(n^{-\mu n})$. Therefore, we only focus on the case when the event $\sum_{i,j} \delta_{ij} > 0$ happens. In the existing literature [14, 27, 18], the instances obeying the Bernoulli model are proven to have no spurious local minima. We show that our complexity metric is able to characterize this property by proving an upper bound on the complexity metric.

Theorem 3.5 *Given $\mu \in [1, n]$ and $p \in (0, 1]$, suppose that the weight matrix C obeys the Bernoulli model with the parameter p and that u^* satisfies the μ-incoherence condition. If $\eta > 2$ is a constant and the sampling rate satisfies*

$$p \geq \min\left\{1, \frac{16(1 + \eta\mu) \log n + 16}{n}\right\},$$

it holds with probability at least $1 - 3n^{-\eta/2+1}$ that

$$\mathbb{D}_\alpha(C, u^*) \leq \max\left\{\frac{3n}{4\alpha}, \frac{1}{2(1 - \alpha)\mu}\right\} \cdot \left(\frac{1}{\mu} - \frac{1}{n}\right)^{-1}.$$

We first establish several lemmas before providing the proof of Theorem 3.5. The first lemma is the Chernoff bound for the sum of Bernoulli random variables [51].

Lemma 3.4 *Suppose that $X_1, \ldots, X_m$ are i.i.d. Bernoulli random variables with the parameter p. Then, it holds that*

$$\mathbb{P}\left(\sum_{i \in [m]} X_i \leq \frac{mp}{2}\right) \leq \exp\left(\frac{-mp}{8}\right), \quad \mathbb{P}\left(\sum_{i \in [m]} X_i \geq \frac{3mp}{2}\right) \leq \exp\left(\frac{-mp}{10}\right).$$

The next lemma provides an upper bound on the total number of nonzero entries.

Lemma 3.5 *Suppose that $n \geq 3$. With probability at least $1 - \exp(-np/10)$, there are at most $3n^2p/2$ nonzero entries in C. With the same probability, it holds that*

$$C_{ij} \geq \frac{2}{3n^2p}, \quad \forall i, j \in [n] \quad \text{s.t.} \quad C_{ij} > 0.$$

Proof For the $n(n-1)$ non-diagonal entries of C, Lemma 3.4 implies that there are at most $(3/2) \cdot n(n-1)p$ nonzero entries with probability at least $1 - \exp\left(-n(n-1)p/20\right)$. For the n diagonal entries of C, the same lemma implies that there are at most $(3/2) \cdot np$ nonzero entries with probability at least $1 - \exp\left(-np/10\right)$. Combining both parts concludes that there are at most $(3/2) \cdot n^2p$ nonzero entries in C with probability at least

$$1 - \exp\left(-n(n-1)p/20\right) - \exp\left(-np/10\right) \geq 1 - 2\exp\left(-np/10\right),$$

where the last inequality is due to $n \geq 3$. The lower bound on C_{ij} follows from the normalization constraint. $\qquad\qquad\square$

For every fixed global solution $\tilde{u}^*$, the next lemma estimates the distance between $(C, \tilde{u}^*)$ and $\mathcal{D}$.

Lemma 3.6 *Suppose that $\tilde{u}^*$ is a given vector and the random matrix C obeys the Bernoulli model. In addition, suppose that $\eta > 2$ is a constant and*

$$\|\tilde{u}^*\|_0 \geq \frac{n}{2\mu}, \quad p \geq \min\left\{1, \frac{16(1 + \eta\mu)\log n + 16}{n}\right\},$$

where $\|\tilde{u}^*\|_0$ is the number of nonzero entries of $\tilde{u}^*$. For every instance $(\tilde{C}, \tilde{u}^*) \in \overline{\mathcal{D}}$, it holds with probability at least $1 - 3n^{-\eta/2}$ that

$$\|C - \tilde{C}\|_1 \geq \frac{4(\|\tilde{u}^*\|_0 - 1)}{3n^2}.$$

Proof For all $i, j \in [n]$, we define Bernoulli random variables X_{ij} to be 1 if $C_{ij} > 0$ and 0 otherwise. Then, X_{ij} are independent identically distributed Bernoulli random variables with the parameter p. Let $N := \sum_{i,j} X_{ij}$ be the number of nonzero weights in C. By the definition of the Bernoulli model, all nonzero entries of C are equal to N^{-1}. Since the global solution $\tilde{u}^*$ is fixed, we assume without loss of generality that

$$\mathcal{I}_1(C, \tilde{u}^*) = [\ell], \quad \mathcal{I}_0(C, \tilde{u}^*) = \{\ell + 1, \ldots, n\}.$$

We fix $\tilde{C}$ to be a weight matrix such that $(\tilde{C}, \tilde{u}^*) \in \overline{\mathcal{D}}$ and investigate three cases.

Case I. We first consider the case when $\mathbb{G}_1(\tilde{C}, \tilde{u}^*)$ is disconnected. Suppose that $\tilde{\mathcal{I}}_{11}$ and $\tilde{\mathcal{I}}_{12}$ are a division of $[\ell]$ such that the nodes in $\tilde{\mathcal{I}}_{11}$ are not connected with the nodes in $\tilde{\mathcal{I}}_{12}$. In addition, we denote $k := |\tilde{\mathcal{I}}_{11}|$ and assume that $k \leq \ell/2$. Since the nodes in $\tilde{\mathcal{I}}_{11}$ are disconnected from the nodes in $\tilde{\mathcal{I}}_{12}$, at least

$$2\sum\nolimits_{i\in\tilde{\mathcal{I}}_{11}, j\in\tilde{\mathcal{I}}_{12}} X_{ij}$$

nonzero entries in C are equal to 0 in $\tilde{C}$. Therefore, we have

$$\|C - \tilde{C}\|_1 \geq \frac{1}{N} \cdot 4 \sum_{i\in\tilde{\mathcal{I}}_{11}, j\in\tilde{\mathcal{I}}_{12}} X_{ij} = \frac{4}{N} \sum_{i\in\tilde{\mathcal{I}}_{11}, j\in\tilde{\mathcal{I}}_{12}} X_{ij}.$$

Using Lemma 3.4, it holds that

$$\sum_{i\in\tilde{\mathcal{I}}_{11}, j\in\tilde{\mathcal{I}}_{12}} X_{ij} \geq \frac{1}{2} \cdot |\tilde{\mathcal{I}}_{11}||\tilde{\mathcal{I}}_{12}|p = \frac{k(\ell - k)p}{2}$$

with probability at least $1 - \exp(-k(\ell - k)p/8)$. Since $k(\ell - k) \geq \ell - 1$, one can write:

$$(3.9) \qquad \|C - \tilde{C}\|_1 \geq \frac{4}{N} \sum_{i\in\tilde{\mathcal{I}}_{11}, j\in\tilde{\mathcal{I}}_{12}} X_{ij} \geq \frac{4}{N} \cdot \frac{(\ell - 1)p}{2} = \frac{2(\ell - 1)p}{N}$$

with the same probability. Considering the union bound over all weight matrices $\tilde{C}$ for which $\mathbb{G}_1(\tilde{C}, \tilde{u}^*)$ is disconnected, inequality (3.9) holds with probability at least

$$1 - \sum_{k=1}^{\lfloor \ell/2 \rfloor} \binom{\ell}{k} \exp\left[-\frac{k(\ell - k)p}{8}\right] \geq 1 - \sum_{k=1}^{\lfloor \ell/2 \rfloor} \left(\frac{\ell e}{k}\right)^k \exp\left[-\frac{k(\ell - k)p}{8}\right]$$

$$= 1 - \sum_{k=1}^{\lfloor \ell/2 \rfloor} \exp\left[k + k\log\left(\frac{\ell}{k}\right) - \frac{k(\ell - k)p}{8}\right],$$

where the inequality uses the relation $\binom{\ell}{k} \leq (\ell e/k)^k$. Using the relation that $k \leq \ell/2$, we can estimate that

$$\exp\left[k + k\log\left(\frac{\ell}{k}\right) - \frac{k(\ell - k)p}{8}\right] \leq \exp\left[k + k\log\ell - \frac{k\ell p}{16}\right]$$

$$= \exp\left[-\frac{k\ell}{16}\left(p - \frac{16(1 + \log\ell)}{\ell}\right)\right] \leq \exp\left[-\frac{k\ell}{16}\left(p - \frac{16(1 + \log n)}{n}\right)\right]$$

$$\leq \exp\left[-\frac{k\ell}{16} \cdot \frac{16\eta\mu\log n}{n}\right] = \exp\left(-\frac{\eta\mu k\ell \log n}{n}\right) = n^{-\frac{\eta\mu\ell}{n} \cdot k} \leq n^{-\frac{\eta}{2} \cdot k},$$

where the second last inequality is from the assumption on p and the last inequality is from $\ell \geq n/(2\mu)$. By taking the summation over $k = 1, \ldots, \lfloor \ell/2 \rfloor$, it follows that

$$1 - \sum_{k=1}^{\lfloor \ell/2 \rfloor} \exp\left[k + k\log\left(\frac{\ell}{k}\right) - \frac{k(\ell - k)p}{8}\right] \geq 1 - \sum_{k=1}^{\lfloor \ell/2 \rfloor} n^{-\frac{\eta}{2} \cdot k}$$

$$\geq 1 - \frac{n^{-\frac{\eta}{2}}}{1 - n^{-\frac{\eta}{2}}} \geq 1 - 2n^{-\eta/2},$$

where the last inequality is due to $n^{-\eta/2} \geq n^{-1} \geq 1/2$. Therefore, inequality (3.9) holds with probability at least $1 - 2n^{-\eta/2}$. Using the lower bound of N in Lemma 3.5, the distance between C and $\tilde{C}$ is at least

$$\frac{2}{3n^2 p} \cdot 2(\ell - 1)p = \frac{4(\ell - 1)}{3n^2}$$

with probability at least

$$1 - 2n^{-\eta/2} - \exp(-np/10) \geq 1 - 2n^{-\eta/2} - n^{-4\mu\eta/5} \geq 1 - 3n^{-\eta/2}.$$

Case II. For the case when $\mathcal{I}_{00}(\tilde{C}, \tilde{u}^*)$ is non-empty, the analysis is the same as *Case I.* and it holds that

$$\|C - \tilde{C}\|_1 \geq \frac{2}{3n^2 p} \cdot 2(\ell - 1)p = \frac{4(\ell - 1)}{3n^2}$$

with probability at least $1 - 3n^{-\eta/2}$.

Case III. Finally, we consider the case when $\mathbb{G}_1(\tilde{C}, \tilde{u}^*)$ is bipartite. In this case, we show that there exists a set of indices $\mathcal{I} \subset [n]^2$ with at least $\max\{\ell^2/2, 1\}$ elements such that
$$\tilde{C}_{ij} = 0, \quad \forall (i, j) \in \mathcal{I}.$$
The proof of the above claim can be found in the proof of Theorem 3.3 and we omit it here. If $\ell \geq 2$, we have $\ell^2/2 \geq 2(\ell - 1)$ and the proof is the same as *Case I.* Otherwise if $\ell = 1$, the inequality

$$\|C - \tilde{C}\|_1 \geq \frac{4(\ell - 1)}{3n^2} = 0$$

always holds.

By combining the above three cases, it holds with probability at least $1 - 9n^{-\eta/2}$ that

$$\|C - \tilde{C}\|_1 \geq \frac{4(\ell - 1)}{3n^2}.$$

$\square$

Now, we are ready to prove Theorem 3.5.

Proof (Proof of Theorem 3.5) Suppose that the instance $\mathcal{MC}(\tilde{C}, \tilde{u}^*) \in \overline{\mathcal{D}}$ attains the maximum in (2.2). Denote

$$d_\alpha := \alpha \|C - \tilde{C}\|_1 + (1 - \alpha)\|u^* - \tilde{u}^*\|_1.$$

Let

$$k := |\mathcal{I}_1(C, u^*)|, \quad \ell := |\mathcal{I}_1(\tilde{C}, \tilde{u}^*)|.$$

Similar to Theorem 3.3, our goal is to decide the optimal global solution $\tilde{u}^*$. By Lemma 3.6, the high-probability lower bound of $\|C - \tilde{C}\|_1$ is increasing in ℓ. Hence, the optimal choice of ℓ is not larger than k. We then analyze two cases.

Case I. We first consider the case when $\ell < n/(2\mu)$. Since $\ell \geq 1$, it follows that $\mu < n/2$. By Lemma 3.2, at least $k - \ell > n/(2\mu)$ nonzero entries in u^* are equal to 0 in $\tilde{u}^*$. Hence, the distance between u^* and $\tilde{u}^*$ satisfies

$$\|u^* - \tilde{u}^*\|_1 \geq 2 \left(1 - \frac{n}{2\mu} \cdot \frac{\mu}{n} \right) \geq 1.$$

Therefore, it holds that

$$\mathbb{D}_\alpha(C, u^*) = d_\alpha^{-1} = \left[\alpha \|C - \tilde{C}\|_1 + (1-\alpha) \|u^* - \tilde{u}^*\|_1 \right]^{-1}$$
$$\leq \frac{1}{1-\alpha} \leq \frac{1}{2(1-\alpha)} \cdot \left(1 - \frac{\mu}{n} \right)^{-1} = \frac{1}{2(1-\alpha)\mu} \cdot \left(\frac{1}{\mu} - \frac{1}{n} \right)^{-1}.$$

Case II. Next, we focus on the case when $\ell \geq n/(2\mu)$. By Lemma 3.6, it holds with probability at least $1 - 3n^{-\eta/2}$ that

$$(3.10) \qquad\qquad \|C - \tilde{C}\|_1 \geq \frac{4(\ell - 1)}{3n^2}.$$

By considering the union bound over $\ell \in \mathcal{L} := \{\lceil n/(2\mu) \rceil, \ldots, k\}$, the probability that inequality (3.10) holds for all $\ell \in \mathcal{L}$ is at least

$$1 - \left(\ell - \frac{n}{2\mu} \right) \cdot 3n^{-\eta/2} \geq 1 - 3n^{-\eta/2+1}.$$

In the remainder of this proof, we assume that inequality (3.10) holds for all $\ell \in \mathcal{L}$. In addition, we assume without loss of generality that

$$|u_1^*| \geq |u_2^*| \geq \cdots \geq |u_k^*| > 0; \quad |u_i^*| = 0, \quad \forall i \geq k + 1.$$

By the assumption of this case, at least $k - \ell$ nonzero entries in u^* are equal to 0 in $\tilde{u}^*$. Then, we can estimate that

$$d_\alpha \geq \min_{n/(2\mu) \leq \ell \leq k} \left[\frac{4\alpha(\ell-1)}{3n^2} + 2(1-\alpha) \sum_{i=\ell+1}^{k} |u_i^*| \right]$$
$$\geq \min_{1 \leq \ell \leq k} \left[\frac{4\alpha(k-1)}{3n^2} + 2(1-\alpha) \sum_{i=\ell+1}^{k} |u_i^*| \right].$$

The above minimization problem can be solved in closed form, which leads to

$$d_\alpha \geq \sum\nolimits_{\ell=1}^{k} \min \left\{ \frac{4\alpha}{3n^2}, 2(1-\alpha)|u_i^*| \right\}.$$

By the second property in Lemma 3.2, we have

$$d_\alpha \geq \sum_{i=2}^{k} \min \left\{ \frac{4\alpha}{3\mu n} |u_i^*|, 2(1-\alpha)|u_i^*| \right\} = \min \left\{ \frac{4\alpha}{3\mu n}, 2(1-\alpha) \right\} \sum_{i=2}^{k} |u_i^*|$$
$$\geq \min \left\{ \frac{4\alpha}{3\mu n}, 2(1-\alpha) \right\} \cdot \left(1 - \frac{\mu}{n} \right) = \min \left\{ \frac{4\alpha}{3n}, 2(1-\alpha)\mu \right\} \cdot \left(\frac{1}{\mu} - \frac{1}{n} \right).$$

The desired upper bound follows from $\mathbb{D}_\alpha(C, u^*) = d_\alpha^{-1}$.

By combining the above two cases, the distance d_α satisfies

$$\mathbb{D}_\alpha(C, u^*) \leq \max\left\{ \frac{3n}{4\alpha}, \frac{1}{2(1-\alpha)\mu} \right\} \cdot \left(\frac{1}{\mu} - \frac{1}{n} \right)^{-1}$$

with probability at least $1 - 3n^{-\eta/2+1}$. □

By Theorem 3.5, if $1 - \alpha = \Theta(n^{-1}\mu^{-1})$, then the complexity of instances obeying the Bernoulli model is on the order of $\Theta[n^2\mu/(n-\mu)]$. If the incoherence $\mu = O(1)$, the complexity is on the order of $O(n)$, which matches the minimum possible complexity up to a constant. Therefore, the proposed metric can also serve as a good indicator for the matrix completion problem with the Bernoulli model. Finally, we note that the bound $p \geq \Theta(\mu \log n/n)$ is optimal up to a constant [15]; see also the discussions in Appendix E of [25].

3.3 One-parameter class of instances

In Sections 3.1 and 3.2, we provided several upper bounds on the complexity metric. In this part, we consider a class of instances that are parameterized by a single parameter $\epsilon \in [0, 1]$. Intuitively, when the parameter grows from 0 to 1, the optimization landscape of the instance becomes more benign. Unlike the previous results in this section, the analysis of the small parameter case provides necessary conditions for the existence of spurious local minima. More specifically, we fix $\mathbb{G} = (\mathbb{V}, \mathbb{E})$ to be an unweighted undirected graph without self-loops, where the node set is $\mathbb{V} = [n]$. We consider the maximal independent set of $\mathbb{G}$, which is defined as follows:

Definition 3.3 For an undirected graph $\mathbb{G} = (\mathbb{V}, \mathbb{E})$, a set $\mathcal{S} \subset \mathbb{V}$ is called an *independent set* if no two nodes in $\mathcal{S}$ are adjacent. The set $\mathcal{S}$ is called a *maximal independent set* if it is an independent set with the maximum number of nodes [2].

Suppose that $\mathcal{S} \subset [n]$ is a maximal independent set of $\mathbb{G}$. For every $\epsilon \in [0, 1]$, the instance $\mathcal{MC}(C^\epsilon, u^*)$ is defined by

$$(3.11) \quad C_{ij}^\epsilon := \epsilon/Z_\epsilon, \quad \forall i, j \in \mathcal{S} \quad \text{s.t.} \quad i \neq j; \quad C_{ij}^\epsilon := 1/Z_\epsilon, \quad \text{if } \{i, j\} \in \mathbb{E};$$

$$C_{ii}^\epsilon := 1/Z_\epsilon, \quad \forall i \in [n], \quad C_{ij}^\epsilon := 0, \quad \text{otherwise},$$

$$u_i^* := 1/m, \quad \forall i \in \mathcal{S}; \quad u_i^* := 0, \quad \forall i \notin \mathcal{S},$$

where $m := |\mathcal{S}|$ and $Z_\epsilon := 2|\mathbb{E}| + n + m(m-1)\epsilon$ is the normalization constant. In the remainder of this subsection, we assume without loss of generality that $\mathcal{S} = [m]$.

First, we study for what values of ϵ the instance $\mathcal{MC}(C^\epsilon, u^*)$ has benign landscape or has spurious local minima. We first prove that the *spurious second-order*

---

[2] We note that this definition is different from the common definition of maximum independent set, which only requires that a maximum independent set is not a proper subset of an independent set.

critical points[3] (SSCPs) of the instance $\mathcal{MC}(C^\epsilon, u^*)$ are closely related to those of the m-dimensional problem

$$(3.12) \qquad \min_{x \in \mathbb{R}^m} \quad \sum_{i \in [m]} (x_i^2 - 1)^2 + \epsilon \sum_{i,j \in [m], i \neq j} (x_i x_j - 1)^2.$$

Lemma 3.7 *If problem* (3.12) *has no SSCPs, then the instance $\mathcal{MC}(C^\epsilon, u^*)$ has no SSCPs. In addition, given a number $N \in \mathbb{N}$, suppose that problem (3.12) has N SSCPs with nonzero components at which the objective function has a positive definite Hessian matrix. Then, the instance $\mathcal{MC}(C^\epsilon, u^*)$ has at least N spurious local minima.*

Proof To prove the first part of the theorem, we assume that problem (3.12) has no SSCPs. Suppose that $u^0 \in \mathbb{R}^n$ is a second-order critical point of the instance $\mathcal{MC}(C^\epsilon, u^*)$. Calculating the gradient of $g(u; C, u^*)$ with respect to u_i for any index $i \geq m$ leads to

$$Z_\epsilon \nabla_i g(u^0; C^\epsilon, u^*) = 4(u_i^0)^3 + 4 \sum_{j \in [n], \{i,j\} \in \mathbb{E}} u_i^0 (u_j^0)^2 = 0,$$

where $\nabla_i g(\cdot; C^\epsilon, u^*)$ is i-th component of the gradient. By multiplying u_i^0 on both sides, it follows that

$$4(u_i^0)^4 + 4(u_i^0)^2 \sum_{j \in [n], \{i,j\} \in \mathbb{E}} (u_j^0)^2 = 0,$$

which implies that $u_i^0 = 0$ for all $i \in \{m+1, \ldots, n\}$. Calculating the gradient and the Hessian matrix with respect to $u_{1:m}$ yields that

$$Z_\epsilon \nabla_i g(u^0; C^\epsilon, u^*) = 4\epsilon \sum_{j \in [m], j \neq i} u_j^0 (u_i^0 u_j^0 - 1/m^2)$$
$$+ 4u_i^0 [(u_i^0)^2 - 1/m^2], \quad \forall i \in [m];$$
$$Z_\epsilon \nabla_{ii}^2 g(u^0; C^\epsilon, u^*) = 12(u_i^0)^2 - 4/m^2 + 4\epsilon \sum_{j \in [m], j \neq i} (u_j^0)^2, \quad \forall i \in [m];$$
$$Z_\epsilon \nabla_{ij}^2 g(u^0; C^\epsilon, u^*) = 4\epsilon(2u_i^0 u_j^0 - 1), \quad \forall i, j \in [m] \text{ s.t. } i \neq j,$$

where $\nabla_{ij} g(\cdot; C^\epsilon, u^*)$ is the (i,j)-th component of the Hessian matrix. By defining $x^0 \in \mathbb{R}^m$ as $x_i^0 := m u_i^0$ for all $i \in [m]$, the above gradient and Hessian matrix turn out to be the same as those of problem (3.12). Since the first m entries of $\nabla g(u^0; C^\epsilon, u^*)$ are 0 and the first m-by-m principle sub-matrix of $\nabla^2 g(u^0; C^\epsilon, u^*)$ is positive semi-definite, the point x^0 is a second-order critical point of problem (3.12). In addition, the point u^0 is a global optimum if and only if $|u_i^0| = 1/m$ for all $i \in [m]$, which is further equivalent to $x_i^0 = 1$ for all $i \in [m]$ and x^0 is the global solution to problem (3.12). Therefore, the point x^0 is a SSCP if u^0 is a SSCP, which is a contradiction to the assumption that problem (3.12) has no SSCPs. Therefore, the point u^0 is a global minimum of the instance $\mathcal{MC}(C^\epsilon, u^*)$.

---

[3] A point $u \in \mathbb{R}^n$ is called a spurious second-order critical point if it satisfies the first-order and the second-order necessary optimality conditions and $uu^T \neq u^*(u^*)^T$.

For the second part of the theorem, suppose that x^0 is a SSCP of problem (3.12), where the Hessian matrix is positive definite and $x_i^0 \neq 0$ for all $i \in [m]$. We construct $u^0 \in \mathbb{R}^n$ by setting $u_i^0 := m^{-1}x_i^0$ for all $i \in [m]$ and $u_i^0 = 0$ for all $i \in \{m + 1, \dots, n\}$. By similar calculations, we can prove that the Hessian matrix at u^0 is a block diagonal matrix with two blocks, where the first block is $H(x; \epsilon)$ and the second block is a diagonal matrix with positive diagonal entries. Moreover, the gradient at u^0 is equal to 0. Hence, u^0 is a SSCP with a positive definite Hessian matrix. The construction shows that the mapping from x^0 to u^0 is injective. $\square$

To simplify the notations in the following proofs, we denote the gradient and the Hessian matrix of the objective function of problem (3.12) by

$$g_i(x; \epsilon) := 4\left[x_i^3 - x_i + \epsilon\sum\nolimits_{j \neq i} x_j(x_i x_j - 1)\right], \quad \forall i \in [m];$$

$$H_{ii}(x; \epsilon) := 4\left[3x_i^2 - 1 + \epsilon\sum\nolimits_{j \neq i} x_j^2\right], \quad \forall i \in [m];$$

$$H_{ij}(x; \epsilon) := 4\epsilon(2x_i x_j - 1), \quad \forall i, j \in [m] \quad \text{s.t.} \quad i \neq j.$$

The following theorem guarantees that the instance $\mathcal{MC}(C^\epsilon, u^*)$ does not have spurious local minima when $\epsilon \geq O(m^{-1})$.

Theorem 3.6 *If $\epsilon > 18/m$, the instance $\mathcal{MC}(C^\epsilon, u^*)$ does not have SSCPs, namely, all second-order critical points are global minima associated with the ground truth solution M^*.*

Proof By Lemma 3.7, we only need to prove that problem (3.12) has no SSCPs. The conclusion holds when $\epsilon = 1$ since the δ-RIP$_{2,2}$ condition holds with $\delta = 0$ and the results in [59] guarantee that there is no SSCP. In the remainder of the proof, we assume that $\epsilon \in [0, 1)$. Suppose that $x^0 \in \mathbb{R}^m$ is a second-order critical point of problem (3.12). Denote

$$S_k := \sum\nolimits_{i=1}^m (x_i^0)^k, \quad \forall k \in \mathbb{N}.$$

Using the first-order optimality conditions, we have

$$(3.13) \qquad 0 = \frac{1}{4}\sum\nolimits_{i \in [m]} g_i(x^0; \epsilon) = (1 - \epsilon)S_3 - (1 - \epsilon)S_1 - m\epsilon S_1 + \epsilon S_1 S_2,$$

$$0 = \frac{1}{4}\sum\nolimits_{i \in [m]} x_i^0 g_i(x^0; \epsilon) = (1 - \epsilon)S_4 - (1 - \epsilon)S_2 - \epsilon S_1^2 + \epsilon S_2^2.$$

Using the second-order necessary optimality conditions, the curvatures of the objective function along the directions

$$c_+ := (x_1^0 - 1, \dots, x_m^0 - 1) \quad \text{and} \quad c_- := (x_1^0 + 1, \dots, x_m^0 + 1)$$

are given by

$$c_+^T H(x; \epsilon)c_+/4 = 3(1 - \epsilon)(S_4 - 2S_3 + S_2) + [\epsilon S_2 - (1 - \epsilon)](S_2 - 2S_1 + m)$$
$$+ 2\epsilon(S_2^2 - 2S_2 S_1 + S_1^2) - \epsilon(S_1^2 - 2nS_1 + m^2) \geq 0,$$

$$c_-^T H(x; \epsilon)c_-/4 = 3(1 - \epsilon)(S_4 + 2S_3 + S_2) + [\epsilon S_2 - (1 - \epsilon)](S_2 + 2S_1 + m)$$
$$+ 2\epsilon(S_2^2 + 2S_2 S_1 + S_1^2) - \epsilon(S_1^2 + 2nS_1 + m^2) \geq 0.$$

Using the relations in (3.13), we can write S_3 and S_4 in terms of S_1 and S_2, which leads to

(3.14)
$$[m\epsilon + 5(1 - \epsilon)]S_2 + 4\epsilon S_1^2 - 4[m\epsilon + (1 - \epsilon)] \cdot |S_1| - [m^2\epsilon + m(1 - \epsilon)] \geq 0.$$

Let c be a positive number such that

$$S_1^2 = cS_2.$$

Using Hölder's inequality, we have $c \in [1, m]$. We note that in the case when $S_2 = 0$, it holds that $S_1 = 0$ and we can choose c to be any constant in $[1, m]$. Then, inequality (3.14) can be written as

(3.15)
$$[m\epsilon + 5(1 - \epsilon) + 4\epsilon c]S_2 - 4[m\epsilon + (1 - \epsilon)]\sqrt{c} \cdot \sqrt{S_2} - [m^2\epsilon + m(1 - \epsilon)] \geq 0.$$

Inequality (3.15) is a quadratic inequality in $\sqrt{S_2}$ and thus, it can be solved in closed form, namely, inequality (3.15) is equivalent to

(3.16)
$$\sqrt{S_2}$$
$$\geq \frac{4[m\epsilon + (1 - \epsilon)]\sqrt{c} + \sqrt{4[m\epsilon + (1 - \epsilon)][8m\epsilon c + 4(1 - \epsilon)c + m^2\epsilon + 5m(1 - \epsilon)]}}{2[m\epsilon + 5(1 - \epsilon) + 4\epsilon c]}$$
$$= m\sqrt{m\epsilon + (1 - \epsilon)} \cdot \left[\sqrt{[8m\epsilon + 4(1 - \epsilon)]c + m^2\epsilon + 5m(1 - \epsilon)} \right.$$
$$\left. - \sqrt{4[m\epsilon + (1 - \epsilon)]c} \right]^{-1}.$$

Consider the function

$$e(c) := \sqrt{[8m\epsilon + 4(1 - \epsilon)]c + m^2\epsilon + 5m(1 - \epsilon)} - \sqrt{4[m\epsilon + (1 - \epsilon)]c},$$
$$\forall c \in [1, m],$$

which is the negative of a unimodal function[4]. Hence, the maximum value of $e(c)$ on $[1, m]$ is attained at 1 or m. Let

$$C := m\epsilon > 18.$$

We calculate that

$$e(m) = \sqrt{9m[m\epsilon + (1 - \epsilon)]} - \sqrt{4m[m\epsilon + (1 - \epsilon)]}$$
$$= \sqrt{m[m\epsilon + (1 - \epsilon)]} \leq \sqrt{m(C + 1)} \leq \sqrt{2mC},$$
$$e(1) = \sqrt{8m\epsilon + 4(1 - \epsilon) + m^2\epsilon + 5m(1 - \epsilon)} - \sqrt{4[m\epsilon + (1 - \epsilon)]}$$
$$\leq \sqrt{8C + 4 + mC + 5m} \leq \sqrt{2(m + 8)C}.$$

---

[4] In this work, we say a function $f : \mathbb{R} \mapsto \mathbb{R}$ is a *unimodal function* if there exists a constant $c \in \mathbb{R}$ such that f is increasing on $(-\infty, c]$ and decreasing on $[c, +\infty)$.

Hence, we have

$$e(c) \leq \sqrt{2(m+8)C}, \quad \forall c \in [1, m].$$

By combining with (3.16), it follows that

$$\sqrt{S_2} \geq m\sqrt{C + (1-\epsilon)} \cdot \left[\sqrt{2(m+8)C}\right]^{-1}$$
$$\geq m\sqrt{C} \cdot \left[\sqrt{2(m+8)C}\right]^{-1} = \frac{m}{\sqrt{2(m+8)}},$$

which further leads to

$$(3.17) \qquad\qquad S_2 \geq \frac{m^2}{2(m+8)} \geq \frac{m}{18}.$$

Therefore, we obtain that

$$\frac{\epsilon}{1-\epsilon} S_2 - 1 \geq \frac{\epsilon m}{18} - 1 > 0.$$

Using the first-order optimality condition, each component x_i^0 is the solution to the third-order polynomial equation

$$(3.18) \qquad g_i(x; \epsilon) = x_i^3 + \left[\frac{\epsilon}{1-\epsilon} S_2 - 1\right] x_i - \frac{\epsilon}{1-\epsilon} S_1 = 0, \quad \forall i \in [m].$$

Since the first-order coefficient $\epsilon/[(1-\epsilon)S_2] - 1$ is positive, the derivative of the polynomial is positive and the equation has a unique real root x_0. Hence, we know

$$x_1^0 = \cdots = x_m^0 = x_0.$$

The equation in (3.18) now becomes

$$x_0^3 + \left[\frac{\epsilon}{1-\epsilon} \cdot mx_0^2 - 1\right] x_0 - \frac{\epsilon}{1-\epsilon} \cdot mx_0 = \left[\frac{m\epsilon}{1-\epsilon} + 1\right](x_0^3 - x_0) = 0,$$

which gives $x_0 \in \{-1, 0, 1\}$. If $x_0 \in \{-1, 1\}$, then the point x^0 is a global optimum. Otherwise if $x_0 = 0$, it follows that $x^0 = 0$ and $S_2 = 0$, which contradicts (3.17). Combining the two cases, we conclude that problem (3.12) does not have SSCPs, which implies that the instance $\mathcal{MC}(C^\epsilon, u^*)$ also has no SSCPs. □

Then, we consider the regime of ϵ where the instance $\mathcal{MC}(C^\epsilon, u^*)$ has spurious solutions. The following theorem studies the case when m is an even number.

Theorem 3.7 *Suppose that m is an even number. If $\epsilon < 1/(m+1)$, then the instance $\mathcal{MC}(C^\epsilon, u^*)$ has at least $2^{m/2}$ spurious local minima.*

Proof By Lemma 3.7, we only need to show that problem (3.12) has at least $\binom{m}{m/2}$ SSCPs whose associated Hessian matrices are positive definite and whose components are nonzero. We consider a point $x^0 \in \mathbb{R}^m$ such that

$$(x_i^0)^2 = \frac{1-\epsilon}{1+(m-1)\epsilon} > 0, \quad \forall i \in [m]; \quad \sum_{i \in [m]} x_i^0 = 0.$$

The above equations have a solution since m is an even number. By a direct calculation, we can verify that the gradient $g(x^0; \epsilon)$ is equal to 0. We only need to show that the Hessian matrix $H(x^0; \epsilon)$ is positive definite, namely

$$c^T H(x^0; \epsilon)c > 0, \quad \forall c \in \mathbb{R}^m \backslash \{0\}.$$

The above condition is equivalent to

$$\left[(3+(m-3)\epsilon)\left(x_1^0\right)^2 - 1 + \epsilon \right] \sum_{i \in [m]} c_i^2 - \epsilon \left(\sum_{i \in [m]} c_i \right)^2$$
$$+ 2\epsilon \left(x_1^0\right)^2 \left(\sum_{i \in [m]} \text{sign}(x_i^0)c_i \right)^2 > 0, \quad \forall c \in \mathbb{R}^n \backslash \{0\}.$$

Under the normalization constraint $\|c\|_2 = 1$, the Cauchy inequality implies that the minimum of the left-hand side is attained by

$$c_1 = \cdots = c_m = 1/\sqrt{m}.$$

Therefore, the Hessian is positive definite if and only if

$$(3+(m-3)\epsilon)\left(x_1^0\right)^2 - 1 + \epsilon > m\epsilon.$$

By substituting $(x_1^0)^2 = (1-\epsilon)/[1+(m-1)\epsilon]$, the above condition is equivalent to

$$2 - (m+4)\epsilon - (m-2)(m+1)\epsilon^2 > 0.$$

Using the condition that $(m+1)\epsilon < 1$, we obtain that

$$2 - (m+4)\epsilon - (m-2)(m+1)\epsilon^2 > 1 - 3\epsilon - (m-2)\epsilon = 1 - (m+1)\epsilon > 0,$$

where the first inequality is from the fact that $m \geq 2$, which follows from the assumption that $m > 0$ is an even number.

To estimate the number of SSCPs, we observe that $m/2$ components of x^0 have a positive sign and the other $m/2$ components have a negative sign. Hence, there are at least

$$\binom{m}{m/2}$$

spurious SSCPs. The estimate on the combinatorial number is in light of the inequality $\binom{n}{k} \geq (n/k)^k$. □

The estimation of the odd number case is similar. We present the result in the following theorem and provide the proof in the appendix.
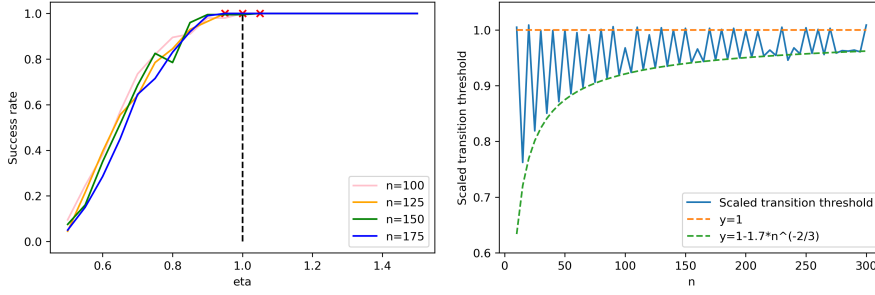
Fig. 2: The left plot shows the transitions of the success rate of the gradient descent algorithm when $n = 100, 125, 150, 175$. The red "$\times$" sign refers to the *transition threshold*, i.e., the smallest value of η that attains 100% success rate. In the right plot, the transition thresholds of η are compared with the curves $y = 1$ and $y = 1 - 1.7(n+1)^{-2/3}$.

Theorem 3.8 *Suppose that m is an odd number. If $\epsilon < 1/[13(m+1)]$, then the instance $\mathcal{MC}(C^\epsilon, u^*)$ has at least $[2m/(m+1)]^{(m+1)/2}$ spurious local minima.*

By combining Theorems 3.6-3.8, it follows that the threshold $\epsilon = \Theta(m^{-1}) = \Theta(\mu/n)$ separates the regimes where the instance possesses and does not possess spurious local minima, where $\mu := n/m$ is the incoherence of u^*. In addition, in the case when $m = 2$, Theorem 3.7 implies that the instance $\mathcal{MC}(C^\epsilon, u^*)$ has spurious local minima if $\epsilon < 1/3$. By Lemma 3.1, the condition that $\epsilon = 1/3$ corresponds to the condition that the δ-$\mathrm{RIP}_{2,2}$ condition holds with $\delta = 1/2$. Therefore, the RIP constant $\delta \leq 1/2$ is necessary for the instance $\mathcal{MC}(C^\epsilon, u^*)$ to have no spurious local minima. Combined with the results in [57,7], we can see that the one-parameter group $\mathcal{MC}(C^\epsilon, u^*)$ also contains difficult instances of the general problem (1.2).

Furthermore, we note that the constants in Theorems 3.6-3.8 are not optimal. We conjecture that the instance $\mathcal{MC}(C^\epsilon, u^*)$ has spurious solutions if $\epsilon < (m+1)^{-1} + o(m^{-1})$ and does not have spurious solutions if $\epsilon > (m+1)^{-1} + o(m^{-1})$. We numerically verify this conjecture in the special case when $m = n$. In numerical examples, we consider the scaled parameter $\eta := (n+1)\epsilon$. For each instance, we implement the randomly initialized gradient descent algorithm for 200 times and check the number of implements for which the distance between the last iterate and $\pm u^*$ has Frobenius norm at most 10^{-5}. The results are plotted in Figure 2. In the left plot, we can see that in most cases, the success rate grows with the parameter η, which is proportional to ϵ. This indicates that the optimization landscape becomes more benign when ϵ is larger. In addition, the transition thresholds of η are very close to 1 (to be more accurate, the thresholds of η are between 0.95 and 1.05). This observation is consistent with our conjecture. In the right plot, we compare the transition thresholds of η against the constant number 1. We observe that the thresholds are approximately located between 1 and $1 - 1.7(n+1)^{-2/3}$, which implies that the original thresholds of ϵ are between $(n+1)^{-1}$ and $(n+1)^{-1} - 1.7(n+1)^{-5/3}$. Hence, the thresholds become close to $(n+1)^{-1}$ when n is large, which is also consistent with our conjecture. Moreover,

we can see that the threshold of η is not monotone in n and is slightly smaller when n is odd.

Finally, we transform the above estimates on the parameter ϵ to the complexity metric.

Theorem 3.9 *Suppose that $n \geq m \geq 36$, $\alpha \in [0, 1]$ and $\epsilon \in [0, 1]$. Then, the following statements hold true:*

1. *If*

$$\mathbb{D}_\alpha(C^\epsilon, u^*) \leq \left[\frac{36\alpha}{n^2} + \min\left\{ 72\alpha \cdot \frac{m}{n^2}, 2(1-\alpha) \right\} \right]^{-1},$$

   then the instance $\mathcal{MC}(C^\epsilon, u^*)$ has no spurious local minima;
2. *If*

$$\mathbb{D}_\alpha(C^\epsilon, u^*) \geq \frac{18}{17} \max\left\{ \frac{13n^2}{2\alpha}, \frac{1}{2(1-\alpha)} \right\},$$

   then the instance $\mathcal{MC}(C^\epsilon, u^*)$ has spurious local minima.

The proof of Theorem 3.9 relies on the following lemma, which calculates the complexity metric of the instance $\mathcal{MC}(C^\epsilon, u^*)$.

Lemma 3.8 *Suppose that $n \geq m \geq 5$, $\alpha \in [0, 1]$ and $\epsilon \in [0, 1]$. The complexity metric $\mathbb{D}_\alpha(C^\epsilon, u^*)$ has the closed form*

$$[\mathbb{D}_\alpha(C^\epsilon, u^*)]^{-1} = \min\left\{ \frac{2\alpha}{Z_\epsilon} + \frac{2(1-\alpha)(m-1)}{m}, \frac{4\alpha\epsilon}{Z_\epsilon} + \frac{2(1-\alpha)(m-2)}{m}, \right.$$
$$\left. \frac{4\alpha(m-1)\epsilon}{Z_\epsilon} \right\}.$$

Moreover, $\mathbb{D}_\alpha(C^\epsilon, u^*)$ is strictly decreasing in ϵ on $[0, 1/2]$.

The proof of Lemma 3.8 is similar to that of Theorem 2.4 and is provided in the appendix. Combining Theorems 3.6-3.8 and Lemma 3.8, we are able to estimate the range of the complexity metric.

Proof (Proof of Theorem 3.9) By defining constants $\delta := 1/26$ and $\Delta := 18$, Theorems 3.6-3.8 imply that

1. If $\epsilon < \delta/m$, the instance $\mathcal{MC}(C^\epsilon, u^*)$ has spurious local minima;
2. If $\epsilon > \Delta/m$, the instance $\mathcal{MC}(C^\epsilon, u^*)$ has no spurious local minima.

Then, we study two different cases.

Case I. We first consider the case when $m\epsilon$ is large. Since $\epsilon < \Delta/m \leq 1/2$, the threshold locates in the regime where $\mathbb{D}_\alpha(C^\epsilon, u^*)$ is strictly decreasing. Hence, it suffices to show that

$$\left[\frac{2\alpha\Delta}{n^2} + \min\left\{ 4\alpha\Delta \cdot \frac{m}{n^2}, 2(1-\alpha) \right\} \right]^{-1}$$

is a lower bound on $\mathbb{D}_\alpha(C^\epsilon, u^*)$ when $\epsilon = \Delta/m$. By Lemma 3.8, it holds that

$$
\begin{aligned}
&[\mathbb{D}_\alpha(C^\epsilon, u^*)]^{-1} \\
&= \min\left\{ \frac{2\alpha}{Z_\epsilon} + \frac{2(1-\alpha)(m-1)}{m}, \frac{4\alpha\epsilon}{Z_\epsilon} + \frac{2(1-\alpha)(m-2)}{m}, \frac{4\alpha(m-1)\epsilon}{Z_\epsilon} \right\} \\
&\leq \min\left\{ \frac{4\alpha\epsilon}{Z_\epsilon} + \frac{2(1-\alpha)(m-2)}{m}, \frac{4\alpha(m-1)\epsilon}{Z_\epsilon} \right\} \\
&= \frac{4\alpha\epsilon}{Z_\epsilon} + (m-2)\min\left\{ \frac{4\alpha\epsilon}{Z_\epsilon}, \frac{2(1-\alpha)}{m} \right\} \leq \frac{4\alpha\epsilon}{Z_\epsilon} + m\min\left\{ \frac{4\alpha\epsilon}{Z_\epsilon}, \frac{2(1-\alpha)}{m} \right\}.
\end{aligned}
$$

Since the graph $\mathbb{G}$ does not contain any independence set with $m+1$ nodes, Turán's theorem [3] implies that the graph $\mathbb{G}$ has at least $n^2/(2m)$ edges, namely,

$$
|\mathbb{E}| \geq n^2/(2m).
$$

We note that the above bound is asymptotically tight and is attained by the Turán graph. Hence, we obtain that

$$
Z_\epsilon = 2|\mathbb{E}| + n + m(m-1)\epsilon \geq 2|\mathbb{E}| \geq n^2/m.
$$

By substituting into the estimate of $\mathbb{D}_\alpha(C^\epsilon, u^*)$, it follows that

$$
\begin{aligned}
[\mathbb{D}_\alpha(C^\epsilon, u^*)]^{-1} &\leq \frac{4\alpha\epsilon \cdot m}{n^2} + m\min\left\{ \frac{4\alpha\epsilon \cdot m}{n^2}, \frac{2(1-\alpha)}{m} \right\} \\
&= \frac{2\alpha\Delta}{n^2} + \min\left\{ 4\alpha\Delta \cdot \frac{m}{n^2}, 2(1-\alpha) \right\}.
\end{aligned}
$$

Case II. Next, we consider the case when ϵm is small. Similar to *Case I*, it suffices to show that

$$
\frac{18}{17} \max\left\{ \frac{n^2}{4\alpha\delta}, \frac{1}{2(1-\alpha)} \right\}
$$

is an upper bound for $\mathbb{D}_\alpha(C^\epsilon, u^*)$ when $\epsilon = \delta/m$. Since $\delta < 1/2$, we have

$$
2\alpha/Z_\epsilon > 4\alpha\epsilon/Z_\epsilon.
$$

By Lemma 3.8, it holds that

$$
\begin{aligned}
&[\mathbb{D}_\alpha(C^\epsilon, u^*)]^{-1} \\
&= \min\left\{ \frac{2\alpha}{Z_\epsilon} + \frac{2(1-\alpha)(m-1)}{m}, \frac{4\alpha\epsilon}{Z_\epsilon} + \frac{2(1-\alpha)(m-2)}{m}, \frac{4\alpha(m-1)\epsilon}{Z_\epsilon} \right\} \\
&= \min\left\{ \frac{4\alpha\epsilon}{Z_\epsilon} + \frac{2(1-\alpha)(m-2)}{m}, \frac{4\alpha(m-1)\epsilon}{Z_\epsilon} \right\} \\
&= \frac{4\alpha\epsilon}{Z_\epsilon} + (m-2)\min\left\{ \frac{4\alpha\epsilon}{Z_\epsilon}, \frac{2(1-\alpha)}{m} \right\} \geq \frac{17}{18}\min\left\{ \frac{4\alpha\epsilon m}{Z_\epsilon}, 2(1-\alpha) \right\},
\end{aligned}
$$

where the last inequality is from $m \geq 36$. Since $\epsilon \leq 1$, the definition of Z_ϵ implies that $Z_\epsilon \leq n^2$. By substituting into the estimate of $\mathbb{D}_\alpha(C^\epsilon, u^*)$, it follows that

$$[\mathbb{D}_\alpha(C^\epsilon, u^*)]^{-1} \geq \frac{17}{18} \min\left\{\frac{4\alpha\epsilon m}{n^2}, 2(1-\alpha)\right\} = \frac{17}{18}\min\left\{\frac{4\alpha\delta}{n^2}, 2(1-\alpha)\right\}.$$

By combining *Cases I* and *II*, we complete the proof. □

In the case when $1 - \alpha \geq \Theta(m/n^2)$, the upper bound on $\mathbb{D}_\alpha(C^\epsilon, u^*)$ is on the order of $O(n\mu/\alpha)$, where $\mu := n/m$ is the incoherence of u^*. This result is consistent with the upper bounds in Sections 3.1 and 3.2. On the other hand, the lower bound in Theorem 3.9 is on the order of $O(n^2/\alpha)$ in the case when $1 - \alpha \geq \Theta(n^{-2})$. In Section 4, we make some conjectures based on these observations and provide a partial theoretical explanation.

4 Theoretical results

In this section, we provide a theoretical analysis for the proposed complexity metric (2.2) on the general problem (1.7). Intuitively, we expect the problem (1.7) to have a benign landscape when the complexity metric is small and vice versa. Recalling the analysis in Section 3, we have the following two conjectures:

Conjecture 1 Suppose that $1 - \alpha \geq \Theta(n^{-1})$ and the solution u^* is μ-incoherent. Then, there exist two constants $\delta, \Delta > 0$ such that

1. If $\mathbb{D}_\alpha(C, u^*) \leq \delta\mu n/\alpha$, the instance $\mathcal{MC}(C, u^*)$ has no SSCPs;
2. If $\mathbb{D}_\alpha(C, u^*) \geq \Delta n^2/\alpha$, the instance $\mathcal{MC}(C, u^*)$ has SSCPs.

Suppose that the first conjecture holds. The results in Section 3.1 imply that the proposed complexity metric guarantees the in-existence of SSCPs when the RIP constant is $O[(\delta - 1)/(\delta + 1)]$, which is independent of μ. In addition, the matrix completion problem under the Bernoulli model does not have SSCPs when $p \geq O(\mu \log n/n)$, which matches the lower bound in [15]. In Section 4.1, we prove a weaker version of the first conjecture in the case when α is equal to α^* or $\alpha^\diamond$, which are defined in Section 2. We note that both α^* and $\alpha^\diamond$ satisfy the condition that $1 - \alpha = \Theta(n^{-1})$. On the other hand, in Section 4.2, we refute the second conjecture by constructing counterexamples. This observation implies that similar to the RIP constant and the incoherence, the proposed complexity metric cannot provide necessary conditions on the in-existence of spurious local solutions. However, if we substitute the degenerate set $\mathcal{D}$ with a slightly smaller set, we prove that the complexity metric is able to provide a necessary condition.

4.1 Small complexity case

We first consider instances with a small complexity metric. In the case when α is equal to α^* or $\alpha^\diamond$, we prove that $\mathbb{D}_\alpha(C, u^*) \leq \delta n/\alpha$ serves as a sufficient condition

for the in-existence of SSCPs, where $\delta > 0$ is an absolute constant. Since the incoherence μ is at least 1, the aforementioned condition is weaker than the first property in Conjecture 1. By Theorem 2.4, the minimum possible value of the complexity metric is on the order of $O(n/\alpha)$. In this subsection, we show that the constant δ can be chosen such that $\delta n/\alpha$ is strictly larger than the minimum possible complexity. The following theorem deals with the case when $\alpha = \alpha^*$.

Theorem 4.1 *Suppose that $n \geq 5$ and $\alpha = \alpha^*$. Then, there exists a constant $\delta > 1/4$ such that for every instance $\mathcal{MC}(C, u^*)$ satisfying*

$$\mathbb{D}_{\alpha^*}(C, u^*) \leq \delta n/\alpha^*,$$

the instance $\mathcal{MC}(C, u^*)$ does not have any SSCPs.

Since the minimum possible complexity metric is $n/(4\alpha^*)$, the upper bound in Theorem 4.1 is *non-trivial* in the sense that there exist instances satisfying the inequality. By Theorem 2.5, the minimum complexity metric $n/(4\alpha^*)$ is only attained by instances in $\mathcal{M}$, where

$$\mathcal{M} := \left\{ (C, u^*) \ \middle| \ |u_i^*| = 1/n, \ C_{ii} = 0, \quad \forall i \in [n], \right.$$
$$\left. C_{ij} = \frac{1}{n(n-1)}, \quad \forall i, j \in [n], \ i \neq j \right\}.$$

In the next lemma, we prove the strict-saddle property [45] of the ℓ_1-norm for instances in $\mathcal{M}$, which can be viewed as a robust version of the in-existence of SSCPs.

Lemma 4.1 *Suppose that $n \geq 2$ and $(C^0, u^0) \in \mathcal{M}$. Then, there exist a positive constant η_0 and two positive-valued functions $\beta(\eta)$ and $\gamma(\eta)$ such that for all $\eta \in (0, \eta_0]$ and $u \in \mathbb{R}^n$, at least one of the following properties holds:*

1. $\min\{\|u - u^*\|_1, \|u + u^*\|_1\} \leq \eta$;
2. $\|\nabla g(u; C, u^*)\|_\infty \geq \beta(\eta)$;
3. $\lambda_{min}[\nabla^2 g(u; C, u^*)] \leq -\gamma(\eta)$.

We then show that after a sufficiently small perturbation to any point $(C^0, x^0) \in \mathcal{M}$, the new instance does not have any SSCPs.

Lemma 4.2 *Suppose that $n \geq 3$. There exists a small positive constant ϵ such that for every pair $(C^0, u^0) \in \mathcal{M}$ and $(\tilde{C}, \tilde{u}^*)$ satisfying*

$$\alpha^* \|\tilde{C} - C^0\|_1 + (1 - \alpha^*)\|\tilde{u}^* - u^0\|_1 < \epsilon,$$

the instance $\mathcal{MC}(\tilde{C}, \tilde{u}^*)$ does not have SSCPs.

The proofs of the last two lemmas can be found in the appendix. Now, we prove the existence of a non-trivial upper bound on the metric.

Proof (Proof of Theorem 4.1) Let ϵ be the constant in Lemma 4.2. We consider the compact set

$$\mathcal{C} := \left\{ (C, u^*) \,\middle|\, \|C\|_1 = \|u^*\|_1 = 1, \right.$$

$$\left. \alpha^* \|\tilde{C} - C^0\|_1 + (1 - \alpha^*)\|\tilde{u}^* - u^0\|_1 \geq \epsilon, \quad \forall (C^0, u^0) \in \mathcal{M} \right\}.$$

Since the minimum possible complexity metric $n/(4\alpha^*)$ is only attained by points in $\mathcal{M}$, it holds that

$$\mathbb{D}_{\alpha^*}(\mathcal{C}) := \max_{(C, u^*) \in \mathcal{C}} \mathbb{D}_{\alpha^*}(C, u^*) > n/(4\alpha^*).$$

Therefore, choosing

$$\delta := (\alpha^*/n) \cdot \mathbb{D}_{\alpha^*}(\mathcal{C}) > 1/4,$$

we have

$$\mathbb{D}_{\alpha^*}(C, u^*) \leq \delta n/\alpha^* \implies (C, u^*) \notin \mathcal{C}$$
$$\implies \text{the instance } \mathcal{MC}(C, u^*) \text{ has no SSCPs.}$$

This completes the proof. $\qquad\square$

The case when $\alpha = \alpha^\diamond$ can be analyzed in a similar way. We note that the strict-saddle property of the instances in Theorem 2.6 has been established in [33]. Hence, we present the results in the following theorem and omit the proof.

Theorem 4.2 *Suppose that $n \geq 5$ and $\alpha = \alpha^\diamond$. Then, there exists a constant $\delta > 1/4$ such that for every pair (C, u^*) satisfying*

$$\mathbb{D}_{\alpha^\diamond}(C, u^*) \leq \delta n(n+2)/(n+1),$$

the instance $\mathcal{MC}(C, u^*)$ does not have any SSCPs.

Similar to Theorem 4.1, since the minimum possible complexity metric is attained with $\delta = 1/4$, the upper bound in Theorem 4.2 is non-trivial.

4.2 Large complexity case

In this subsection, we first refute the second property in Conjecture 1 and then refine its statement to make it hold true. We note that the RIP constant and the incoherence cannot provide necessary conditions for the in-existence of SSCPs either. Namely, there exist instances that satisfy the δ-RIP$_{2,2}$ condition with δ as high as 1 which do not have SSCPs. Similarly, in the case when the incoherence of the global solution is n, it is still possible to have an instance of the matrix completion problem without any SSCPs. In other words, although small values for the RIP constant and incoherence guarantee the in-existence of spurious solutions, these notions cannot capture the complexity of the problem since there are low-complexity problems with large values for these parameters. We first show that our new metric suffers from the same shortcoming, but we then propose a simple refinement to address this issue.

Example 1 Suppose that the weight matrix and the ground truth are

$$C^\delta := \frac{1}{1 + 3\delta} \begin{bmatrix} 1 & \delta \\ \delta & \delta \end{bmatrix}, \quad u^* := \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

where $\delta \geq 0$ is a constant. One can verify that $\pm u^*$ are the only local minima to the instance $\mathcal{MC}(C^\delta, u^*)$ for all $\delta > 0$. However, in the case when $\delta = 0$, the instance $\mathcal{MC}(C^0, u^*)$ has the set of global solutions

$$\pm \begin{bmatrix} 1 \\ c \end{bmatrix}, \quad \forall c \in \mathbb{R}.$$

Moreover, we consider the case when both components of u^* are measured, where the instance $\mathcal{MC}(\tilde{C}^\epsilon, \tilde{u}^\epsilon)$ is defined by

$$\tilde{C}^\epsilon := \frac{1}{1 + \epsilon} \begin{bmatrix} 1 & 0 \\ 0 & \epsilon \end{bmatrix}, \quad \tilde{u}^\epsilon := \frac{1}{1 + \epsilon} \begin{bmatrix} 1 \\ \epsilon \end{bmatrix},$$

where ϵ is a positive constant. One can verify that the pair $(\tilde{C}^\epsilon, \tilde{u}^\epsilon)$ belongs to $\mathcal{D}$ for all $\epsilon > 0$. Setting δ and ϵ to be small enough, the instances $\mathcal{MC}(C^\delta, u^*)$ and $\mathcal{MC}(\tilde{C}^\epsilon, \tilde{u}^\epsilon)$ can be arbitrarily close to each other in the sense that

$$\alpha \|C^\delta - \tilde{C}^\epsilon\|_1 + (1 - \alpha)\|u^* - \tilde{u}^\epsilon\|_1 = O(\alpha\delta + \epsilon).$$

Therefore, the complexity metric of $\mathcal{MC}(C^\delta, u^*)$ can be arbitrarily large. This example shows that instances without SSCPs can be arbitrarily close to those in $\mathcal{D}$, which have non-unique global solutions.

Nevertheless, we derive a lower bound on the complexity metric (2.2) by constructing a subset of $\mathcal{D}$, which allows obtaining a necessary condition. Intuitively, if an instance has multiple global minima, these global minima are still locally optimal after a sufficiently small perturbation to the instance. To ensure the "robustness" of the local optimality, we require the positive-definiteness of the Hessian matrix. For each instance $\mathcal{MC}(C, u^*)$, let $\mathbb{G}_{1k}(C, u^*)$ for all $k \in [n_1]$ be the connected components of $\mathbb{G}_1(C, u^*)$, where n_1 is the number of connected components. Moreover, we use $\mathcal{I}_{1k}(C, u^*)$ to denote the node set of $\mathbb{G}_{1k}(C, u^*)$ for all $k \in [n_1]$. We define the following subset of $\mathcal{D}$:

$$\mathcal{SD} := \{(C, u^*) \in \mathcal{D} \mid \mathbb{G}_{1k}(C, u^*) \text{ is non-bipartite for all } k \in [n_1],$$
$$\mathbb{G}_1(C, u^*) \text{ is disconnected}, \quad \mathcal{I}_{00}(C, u^*) = \emptyset\}.$$

The following theorem provides a characterization of the Hessian matrix at global solutions for pairs in $\mathcal{SD}$.

Theorem 4.3 *Suppose that $(C, u^*) \in \mathcal{SD}$. Then, the Hessian matrix is positive definite at all global solutions of the instance $\mathcal{MC}(C, u^*)$.*

The proof of Theorem 4.3 directly follows from the next two lemmas.

Lemma 4.3 *Suppose that $(C, u^*) \in \mathcal{SD}$ and that u^0 is a global solution to $\mathcal{MC}(C, u^*)$. Then, for all $k \in [n_1]$, it holds that $u_i^0 u_j^0 = u_i^* u_j^*$ for all $i, j \in \mathcal{I}_{1k}$. In addition, $u_i^0 = 0$ for all $i \in \mathcal{I}_0(C, u^*)$.*

Proof Denote $M^* := u^*(u^*)^T$. We first consider nodes in $\mathcal{G}_{1k}$ for some $k \in [n_1]$. Since the subgraph is non-bipartite, there exists a cycle with an odd length $2\ell + 1$, which we denote as

$$\{i_1, \ldots, i_{2\ell+1}\}.$$

Then, we have

$$
\begin{aligned}
(u_{i_1}^0)^2 &= \prod_{s=1}^{2\ell+1} (u_{i_s}^0 u_{i_{s+1}}^0)^{(-1)^{s-1}} = \prod_{s=1}^{2\ell+1} (M_{i_s i_{s+1}}^*)^{(-1)^{s-1}} \\
&= \prod_{s=1}^{2\ell+1} (u_{i_s}^* u_{i_{s+1}}^*)^{(-1)^{s-1}} = (u_{i_1}^*)^2,
\end{aligned}
$$

which implies that the conclusion holds for $i = j = i_1$. Using the connectivity of $\mathbb{G}_{1k}(C, u^*)$, we know

$$u_i^0 u_j^0 = u_i^* u_k^*, \quad \forall i, j \in \mathcal{I}_{1k}(C, u^*).$$

Then, we consider nodes in $\mathcal{I}_0(C, u^*)$. Since $\mathcal{I}_{00}(C, u^*)$ is empty, for every node $i \in \mathcal{I}_0(C, u^*)$, there exists another node $j \in \mathcal{I}_1(C, u^*)$ such that $C_{ij} > 0$. Hence, we have

$$u_i^0 = M_{ij}^* / u_j^0 = 0.$$

This completes the proof. $\qquad\square$

The following lemma provides a necessary and sufficient condition for instances with a positive definite Hessian matrix at global solutions, which is stronger than what Theorem 4.3 requires.

Lemma 4.4 *Suppose that $u^0 \in \mathbb{R}^n$ is a global minimizer of the instance $\mathcal{MC}(C, u^*)$ such that the conditions in Lemma 4.3 hold. Then, the Hessian matrix is positive definite at u^0 if and only if*

1. *$\mathbb{G}_{1i}(C, u^*)$ is non-bipartite for all $i \in [n_1]$;*
2. *$\mathcal{I}_{00}(C, u^*) = \emptyset$.*

Proof We first construct counterexamples for the necessity part and then prove the positive definiteness of the Hessian matrix for the sufficiency part.

Necessity. We construct counterexamples by discussing two different cases.

Case I. We first consider the case when there exists $k \in [n_1]$ such that $\mathbb{G}_{1k}(C, u^*)$ is bipartite. Suppose that $\mathbb{G}_{1i}(C, u^*) = \mathbb{G}_{1k1} \cup \mathbb{G}_{1k2}$ is a partition of $\mathbb{G}_{1k}(C, u^*)$. Let the sets $\mathcal{I}_{1k}, \mathcal{I}_{1k1}$ and $\mathcal{I}_{1k2}$ be the node sets of the corresponding graphs. Define $q \in \mathbb{R}^n$ as

$$q_i := u_i^0, \quad \forall i \in \mathcal{I}_{1k1}; \quad q_i := -u_i^0, \quad \forall i \in \mathcal{I}_{1k2}; \quad q_i := 0, \quad \forall i \notin \mathcal{I}_{1k}.$$

Then, the curvature of the Hessian along the direction q is

$$
\begin{aligned}
&\frac{1}{4}[\nabla^2 g(u^0; C, u^*)](q, q) \\
&= \sum_{i \in \mathcal{I}_{1k1}, j \in \mathcal{I}_{1k2}} C_{ij} \left[(u_i^0)^2 q_j^2 + (u_j^0)^2 q_i^2\right] + 2 \sum_{i \in \mathcal{I}_{1k1}, j \in \mathcal{I}_{1k2}} C_{ij}(2u_i^0 u_j^0 - u_i^* u_j^*) q_i q_j \\
&= \sum_{i \in \mathcal{I}_{1k1}, j \in \mathcal{I}_{1k2}} C_{ij} \left[(u_i^0)^2 q_j^2 + (u_j^0)^2 q_i^2\right] + 2 \sum_{i \in \mathcal{I}_{1k1}, j \in \mathcal{I}_{1k2}} C_{ij} u_i^0 u_j^0 q_i q_j \\
&= \sum_{i \in \mathcal{I}_{1k1}, j \in \mathcal{I}_{1k2}} 2 C_{ij}(u_i^0 u_j^0)^2 - 2 \sum_{i \in \mathcal{I}_{1k1}, j \in \mathcal{I}_{1k2}} C_{ij}(u_i^0 u_j^0)^2 = 0.
\end{aligned}
$$

We note that there is no self-loop in $\mathcal{G}_{1k}(C, u^*)$ and, thus, the diagonal entries of the weight matrix are equal to 0. Therefore, the Hessian matrix has a zero curvature along q and is not positive definite.

Case II. We consider the case when $\mathcal{I}_{00}(C, u^*) \neq \emptyset$. Suppose that $k \in \mathcal{I}_{00}(C, u^*)$. Define the vector $q \in \mathbb{R}^n$ as

$$
q_k := 1; \quad q_i := 0, \quad \forall i \neq k.
$$

The curvature of the Hessian along the direction q is

$$
\begin{aligned}
\frac{1}{4}[\nabla^2 g(u^0; C, u^*)](q, q) &= C_{kk}\left[2(u_k^0)^2 - (u_k^*)^2\right] q_k^2 + \sum_{j \in \mathcal{I}_0(C, u^*), j \neq k} C_{kj}(u_j^0)^2 q_k^2 \\
&= C_{kk}(u_k^0)^2 + \sum_{j \in \mathcal{I}_0(C, u^*), j \neq k} C_{kj}(u_j^0)^2 = 0.
\end{aligned}
$$

Therefore, the Hessian matrix is not positive-definite at u^0.

Sufficiency. Next, we consider the sufficiency part, namely, we prove that the Hessian matrix is positive definite under the two conditions stated in the theorem. Suppose that there exists a nonzero vector $q \in \mathbb{R}^n$ such that

$$
[\nabla^2 g(u^0; C, u^*)](q, q) = 0.
$$

Then, after straightforward calculations, we arrive at

$$
\begin{aligned}
u_i^0 q_j + u_j^0 q_i &= 0, \quad \forall i, j \quad \text{s.t.} \quad C_{ij} > 0, \ i \neq j; \\
[2(u_i^0)^2 - (u_i^*)^2] q_i^2 = (u_i^0 q_i)^2 &= 0, \quad \forall i \quad \text{s.t.} \quad C_{ii} > 0.
\end{aligned}
$$

The two conditions can be written compactly as

$$
\text{(4.1)} \qquad u_i^0 q_j + u_j^0 q_i = 0, \quad \forall i, j \quad \text{s.t.} \quad C_{ij} > 0.
$$

Consider the index set $\mathcal{I}_{1k}(C, u^*)$ for some $k \in [n_1]$. The equality (4.1) implies that

$$
\text{(4.2)} \qquad q_i / u_i^0 + q_j / u_j^0 = 0, \quad \forall i, j \in \mathcal{I}_{1k}(C, u^*).
$$

Since the graph $\mathbb{G}_{1k}(C, u^*)$ is non-bipartite, there exists a cycle with an odd length $2\ell + 1$, which we denote as

$$\{i_1, i_2, \ldots, i_{2\ell+1}\}.$$

Denoting $i_{2\ell+2} := i_1$, we can calculate that

$$2\frac{q_{i_1}}{u_{i_1}^0} = \sum_{s=1}^{2\ell+1} (-1)^{s-1} \left(\frac{q_{i_s}}{u_{i_s}^0} + \frac{q_{i_{s+1}}}{u_{i_{s+1}}^0} \right) = 0,$$

which leads to $q_{i_1} = 0$. Using the connectivity of $\mathcal{G}_{1k}$ and the relation (4.2), it follows that

$$q_i = 0, \quad \forall i \in \mathcal{I}_{1k}(C, u^*).$$

Moreover, the same conclusion holds for all $k \in [n_1]$ and, thus, we conclude that

$$q_i = 0, \quad \forall i \in \mathcal{I}_1(C, u^*).$$

Since $\mathcal{I}_{00}(C, u^*) = \emptyset$, for every node $i \in \mathcal{I}_0(C, u^*)$, there exists another node $j \in \mathcal{I}_1(C, u^*)$ such that $C_{ij} > 0$. Considering the relation (4.2), we obtain that

$$q_j = -u_j^0 q_i / u_i^0 = 0.$$

In summary, we have proved that $q_i = 0$ for all $i \in [n]$, which contradicts the assumption that $q \neq 0$. Hence, the Hessian matrix at u^0 is positive definite. $\quad\square$

Using the positive-definiteness of the Hessian matrix, we are able to apply the implicit function theorem to certify the existence of spurious local minima.

Lemma 4.5 *Suppose that $\alpha \in [0, 1]$ and consider a pair $(C, u^*) \in \mathcal{SD}$. Then, there exists a small constant $\delta(C, x^*) > 0$ such that for every instance $\mathcal{MC}(\tilde{C}, \tilde{u}^*)$ satisfying*

$$\alpha\|\tilde{C} - C\|_1 + (1 - \alpha)\|\tilde{u}^* - u^*\|_1 < \delta(C, u^*),$$

the instance $\mathcal{MC}(\tilde{C}, \tilde{u}^*)$ has spurious local minima.

Proof By Theorem 4.3, there exists a global solution u^0 to the instance $\mathcal{MC}(C, u^*)$ such that

$$u^0(u^0)^T \neq u^*(u^*)^T, \quad \nabla^2 g(u^0; C, u^*) \succ 0.$$

Consider the system of equations:

$$\nabla g(u; C, u^*) = 0.$$

Since the Jacobi matrix of $\nabla g(u; C, u^*)$ with respect to u is the Hessian matrix $\nabla^2 g(u; C, u^*)$ and (u^0, C, u^*) is a solution, the implicit function theorem guarantees that there exists a small constant $\delta(C, u^*) > 0$ such that in the neighborhood

$$\mathcal{N} := \left\{ (\tilde{C}, \tilde{u}^*) \mid \alpha\|\tilde{C} - C\|_1 + (1 - \alpha)\|\tilde{u}^* - u^*\|_1 < \delta(C, u^*) \right\},$$

there exists a function $u(\tilde{C}, \tilde{u}^*) : \mathcal{N} \mapsto \mathbb{R}^n$ such that

1. $u(C, u^*) = u^0$;

2. $u(\cdot, \cdot)$ is a continuous function in $\mathcal{N}$;
3. $\nabla g[u(\tilde{C}, \tilde{u}^*); \tilde{C}, \tilde{u}^*] = 0$.

Using the continuity of the Hessian matrix and $u(\cdot, \cdot)$, we can choose $\delta(C, u^*)$ to be small enough such that

$$u(\tilde{C}, \tilde{u}^*)[u(\tilde{C}, \tilde{u}^*)]^T \neq \tilde{u}^*(\tilde{u}^*)^T, \quad \nabla^2 g\left[u(\tilde{C}, \tilde{u}^*); \tilde{C}, \tilde{u}^*\right] \succ 0, \quad \forall(\tilde{C}, \tilde{u}^*) \in \mathcal{N}.$$

Therefore, the point $u(\tilde{C}, \tilde{u}^*)$ is a spurious local minimum of the instance $\mathcal{MC}(\tilde{C}, \tilde{u}^*)$.
□

Lemma 4.5 guarantees the existence of spurious local minima in a neighbour-hood of each instance in $\mathcal{SD}$. The global guarantee can be established by considering closed subsets of $\mathcal{SD}$. For every constant $\epsilon \geq 0$, we consider the closed subset $\mathcal{SD}_\epsilon$, which is defined as

$$\mathcal{SD}_\epsilon := \big\{(C, u^*) \in \mathcal{SD} \mid C_{ij} \in \{0\} \cup [\epsilon, 1], \quad \forall i, j \in [n],$$
$$|u_i^*| \in \{0\} \cup [\epsilon, 1], \quad \forall i \in [n]\big\}.$$

Basically, the extra condition in the definition of $\mathcal{SD}_\epsilon$ requires that the nonzero components of C and u^* be at least ϵ. We can verify that the set $\mathcal{SD}_\epsilon$ is a compact set and for every $\epsilon_n \to 0$, it holds that

$$\lim_{n\to\infty} \cup_{i=1}^n \mathcal{SD}_{\epsilon_i} = \mathcal{SD}_0 = \mathcal{SD}.$$

Now, we define the alternative complexity metric

$$(4.3) \qquad \mathbb{D}_{\alpha,\epsilon}(C, u^*) := \left[\min_{(\tilde{C}, \tilde{u}^*) \in \mathcal{SD}_\epsilon} \alpha \|C - \tilde{C}\|_1 + (1 - \alpha)\|u^* - \tilde{u}^*\|_1\right]^{-1}.$$

Since $\mathcal{SD}_\epsilon$ is a subset of $\mathcal{D}$, it holds that

$$\mathbb{D}_{\alpha,\epsilon}(C, u^*) \leq \mathbb{D}_\alpha(C, u^*).$$

Similar to Theorem 2.3, we can prove the following relation:

$$\overline{\mathcal{SD}} = \{(C, u^*) \mid C \in \mathbb{S}_{+,1}^{n^2-1}, u^* \in \mathbb{S}_1^{n-1}, \mathbb{G}_1(C, u^*) \text{ is disconnected}\}$$
$$\cup \{(C, u^*) \mid C \in \mathbb{S}_{+,1}^{n^2-1}, u^* \in \mathbb{S}_1^{n-1}, \mathcal{I}_{00}(C, u^*) \text{ is not empty}\}.$$

Hence, the closure of $\mathcal{SD}$ is a proper subset of $\overline{\mathcal{D}}$. Combining with the fact that $\mathcal{SD}_\epsilon$ is a subset of $\mathcal{SD}$, the metric $\mathbb{D}_{\alpha,\epsilon}(C, u^*)$ is not equivalent to $\mathbb{D}_\alpha(C, u^*)$. Using the compactness of $\mathcal{SD}_\epsilon$, the following theorem provides a necessary condition for the existence of spurious local minima.

Theorem 4.4 *Suppose that $\epsilon > 0$ is a constant. Then, there exists a large constant $\Delta(\epsilon) > 0$ such that for every instance $\mathcal{MC}(C, u^*)$ satisfying*

$$\mathbb{D}_{\alpha,\epsilon}(C, u^*) \geq \Delta(\epsilon),$$

the instance $\mathcal{MC}(C, u^*)$ has spurious local minima.

Proof For every pair $(C, u^*) \in \mathcal{SD}_\epsilon$, Lemma 4.5 implies that there exists an open neighborhood of (C, u^*) such that the desired properties hold. Now, we consider the union of such open neighborhoods over all points $(C, u^*) \in \mathcal{SD}_\epsilon$, which is an open cover of $\mathcal{SD}_\epsilon$. Using the Heine-Borel covering theorem, there exists an open sub-cover of $\mathcal{SD}_\epsilon$. Therefore, we obtain the existence of $\Delta(\epsilon)$. $\square$

We note that the maximum possible value of $\mathbb{D}_{\alpha, \epsilon}(C, u^*)$ is $+\infty$, which is attained by instances in $\mathcal{SD}_\epsilon$. Therefore, there exist instances satisfying the condition of Theorem 4.4 and the lower bound is *non-trivial*. Using Theorem 4.4, the slightly modified complexity metric is able to provide a necessary condition on the in-existence of SSCPs. This result implies that our complexity metric is much better than the RIP constant and the incoherence that fail to provide necessary conditions.

Finally, we conjecture that the second property in Conjecture 1 holds for any fixed weight matrix. More specifically, we define

$$(4.4) \qquad \mathbb{D}_C(u^*) := \left(\min_{(C, \tilde{u}^*) \in \overline{\mathcal{D}}} \|u^* - \tilde{u}^*\|_1 \right)^{-1}.$$

We have the following conjecture:

Conjecture 2 Suppose that $\epsilon \in [0, 1]$. Then, there exists a large constant $\Gamma(\epsilon) > 0$ such that for every instance $\mathcal{MC}(C, u^*)$ satisfying

$$C_{ij} \in \{0\} \cup [\epsilon, 1], \quad \mathbb{D}_C(u^*) \geq \Gamma(\epsilon),$$

the instance $\mathcal{MC}(C, u^*)$ has spurious local minima.

We note that the metric $\mathbb{D}_C(u^*)$ is equal to 0 if $\mathcal{MC}(C, u^*)$ satisfies the δ-RIP$_{2,2}$ condition with $\delta \in [0, 1)$.

5 Conclusions

In this work, we propose a new complexity metric for an important class of the low-rank matrix optimization problems, which has the potential to unify two existing complexity metrics and is applicable to a much broader set of problems. The proposed complexity metric aims to measure the complexity of the non-convex optimization landscape of each problem and quantifies the likelihood of local search methods in successfully solving each instance of the problem under a random initialization. We focus on the rank-1 generalized matrix completion problem (1.7) to mathematically prove the usefulness of the new metric from three aspects. Namely, we show that the complexity metric has a small value if the instance satisfies the RIP condition or obeys the Bernoulli model. The results in these two scenarios are consistent with the existing results on RIP and incoherence. In addition, we analyze a one-parameter class of instances of the matrix sensing problem to illustrate that the proposed metric captures the true complexity of this class as the parameter varies and has consistent behavior with the aforementioned two scenarios. Finally, we provide strong theoretical results on the proposed complexity metric by showing that a small value for the proposed

complexity metric guarantees the in-existence of spurious solutions, whereas a large value for a slightly modified complexity metric guarantees the existence of spurious solutions. This shows the superiority of this metric over the RIP constant and the incoherence since those notions cannot offer any necessary conditions on having spurious solutions.

References

1. Agarwal, A., Anandkumar, A., Jain, P., Netrapalli, P.: Learning sparsely used overcomplete dictionaries via alternating minimization. SIAM Journal on Optimization **26**(4), 2775–2799 (2016)
2. Ahn, K., Suarez, F.: Riemannian perspective on matrix factorization. arXiv preprint arXiv:2102.00937 (2021)
3. Aigner, M.: Turán's graph theorem. The American Mathematical Monthly **102**(9), 808–816 (1995)
4. Ajayi, T., Mildebrath, D., Kyrillidis, A., Ubaru, S., Kollias, G., Bouchard, K.: Provably convergent acceleration in factored gradient descent with applications in matrix sensing. arXiv preprint arXiv:1806.00534 (2018)
5. Allen-Zhu, Z., Li, Y.: Neon2: Finding local minima via first-order oracles. Advances in Neural Information Processing Systems **31** (2018)
6. Bi, Y., Lavaei, J.: On the absence of spurious local minima in nonlinear low-rank matrix recovery problems. In: International Conference on Artificial Intelligence and Statistics, pp. 379–387. PMLR (2021)
7. Bi, Y., Zhang, H., Lavaei, J.: Local and global linear convergence of general low-rank matrix recovery problems. In: Proceedings of 36th AAAI Conference on Artificial Intelligence (AAAI), Vancouver, Canada, pp. 1–9 (2022)
8. Boumal, N., Voroninski, V., Bandeira, A.: The non-convex burer-monteiro approach works on smooth semidefinite programs. Advances in Neural Information Processing Systems **29** (2016)
9. Boumal, N., Voroninski, V., Bandeira, A.S.: Deterministic guarantees for burer-monteiro factorizations of smooth semidefinite programs. Communications on Pure and Applied Mathematics **73**(3), 581–608 (2020)
10. Burer, S., Monteiro, R.D.: A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. Mathematical Programming **95**(2), 329–357 (2003)
11. Candès, E.J., Li, X., Ma, Y., Wright, J.: Robust principal component analysis? Journal of the ACM (JACM) **58**(3), 1–37 (2011)
12. Candes, E.J., Li, X., Soltanolkotabi, M.: Phase retrieval via wirtinger flow: Theory and algorithms. IEEE Transactions on Information Theory **61**(4), 1985–2007 (2015)
13. Candes, E.J., Plan, Y.: Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements. IEEE Transactions on Information Theory **57**(4), 2342–2359 (2011)
14. Candès, E.J., Recht, B.: Exact matrix completion via convex optimization. Foundations of Computational mathematics **9**(6), 717–772 (2009)
15. Candès, E.J., Tao, T.: The power of convex relaxation: Near-optimal matrix completion. IEEE Transactions on Information Theory **56**(5), 2053–2080 (2010)
16. Cartis, C., Gould, N.I., Toint, P.L.: Adaptive cubic regularisation methods for unconstrained optimization. part i: motivation, convergence and numerical results. Mathematical Programming **127**(2), 245–295 (2011)
17. Chen, J., Li, X.: Model-free nonconvex matrix completion: Local minima analysis and applications in memory-efficient kernel pca. J. Mach. Learn. Res. **20**(142), 1–39 (2019)
18. Chen, J., Liu, D., Li, X.: Nonconvex rectangular matrix completion via gradient descent without $\ell_{2,\infty}$ regularization. IEEE Transactions on Information Theory **66**(9), 5806–5841 (2020)
19. Chen, Y., Chi, Y.: Harnessing structures in big data via guaranteed low-rank matrix estimation: Recent theory and fast algorithms via convex and nonconvex optimization. IEEE Signal Processing Magazine **35**(4), 14–31 (2018)
20. Chen, Y., Chi, Y., Fan, J., Ma, C.: Gradient descent with random initialization: Fast global convergence for nonconvex phase retrieval. Mathematical Programming **176**(1), 5–37 (2019)
21. Chen, Y., Chi, Y., Fan, J., Ma, C., Yan, Y.: Noisy matrix completion: Understanding statistical guarantees for convex relaxation via nonconvex optimization. SIAM journal on optimization **30**(4), 3098–3121 (2020)

22. Chen, Y., Fan, J., Ma, C., Yan, Y.: Bridging convex and nonconvex optimization in robust PCA: Noise, outliers and missing data. The Annals of Statistics **49**(5), 2948–2971 (2021)
23. Chi, Y., Lu, Y.M., Chen, Y.: Nonconvex optimization meets low-rank matrix factorization: An overview. IEEE Transactions on Signal Processing **67**(20), 5239–5269 (2019)
24. Chou, H.H., Gieshoff, C., Maly, J., Rauhut, H.: Gradient descent for deep matrix factorization: Dynamics and implicit bias towards low rank. arXiv preprint arXiv:2011.13772 (2020)
25. Fattahi, S., Sojoudi, S.: Exact guarantees on the absence of spurious local minima for non-negative rank-1 robust principal component analysis. Journal of machine learning research (2020)
26. Ge, R., Jin, C., Zheng, Y.: No spurious local minima in nonconvex low rank problems: A unified geometric analysis. In: International Conference on Machine Learning, pp. 1233–1242. PMLR (2017)
27. Ge, R., Lee, J.D., Ma, T.: Matrix completion has no spurious local minimum. Advances in Neural Information Processing Systems pp. 2981–2989 (2016)
28. Hardt, M.: Understanding alternating minimization for matrix completion. In: 2014 IEEE 55th Annual Symposium on Foundations of Computer Science, pp. 651–660. IEEE (2014)
29. Hardt, M., Wootters, M.: Fast matrix completion without the condition number. In: Conference on learning theory, pp. 638–678. PMLR (2014)
30. Hou, T.Y., Li, Z., Zhang, Z.: Fast global convergence for low-rank matrix recovery via riemannian gradient descent with random initialization. arXiv preprint arXiv:2012.15467 (2020)
31. Jain, P., Meka, R., Dhillon, I.: Guaranteed rank minimization via singular value projection. Advances in Neural Information Processing Systems **23** (2010)
32. Jain, P., Netrapalli, P., Sanghavi, S.: Low-rank matrix completion using alternating minimization. In: Proceedings of the forty-fifth annual ACM symposium on Theory of computing, pp. 665–674 (2013)
33. Jin, C., Ge, R., Netrapalli, P., Kakade, S.M., Jordan, M.I.: How to escape saddle points efficiently. In: International Conference on Machine Learning, pp. 1724–1732. PMLR (2017)
34. Jin, C., Netrapalli, P., Jordan, M.I.: Accelerated gradient descent escapes saddle points faster than gradient descent. In: Conference On Learning Theory, pp. 1042–1085. PMLR (2018)
35. Lee, J.D., Simchowitz, M., Jordan, M.I., Recht, B.: Gradient descent only converges to minimizers. In: Conference on learning theory, pp. 1246–1257. PMLR (2016)
36. Li, Y., Ma, T., Zhang, H.: Algorithmic regularization in over-parameterized matrix sensing and neural networks with quadratic activations. In: Conference On Learning Theory, pp. 2–47. PMLR (2018)
37. Luo, Y., Li, X., Zhang, A.R.: Nonconvex factorization and manifold formulations are almost equivalent in low-rank matrix optimization. arXiv preprint arXiv:2108.01772 (2021)
38. Ma, Z., Bi, Y., Lavaei, J., Sojoudi, S.: Sharp restricted isometry property bounds for low-rank matrix recovery problems with corrupted measurements. arXiv preprint arXiv:2105.08232 (2021)
39. Netrapalli, P., Jain, P., Sanghavi, S.: Phase retrieval using alternating minimization. Advances in Neural Information Processing Systems **26** (2013)
40. Netrapalli, P., UN, N., Sanghavi, S., Anandkumar, A., Jain, P.: Non-convex robust pca. Advances in Neural Information Processing Systems **27** (2014)
41. Pumir, T., Jelassi, S., Boumal, N.: Smoothed analysis of the low-rank approach for smooth semidefinite programs. Advances in Neural Information Processing Systems **31** (2018)
42. Recht, B., Fazel, M., Parrilo, P.A.: Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. SIAM review **52**(3), 471–501 (2010)
43. Stöger, D., Soltanolkotabi, M.: Small random initialization is akin to spectral learning: Optimization and generalization guarantees for overparameterized low-rank matrix reconstruction. Advances in Neural Information Processing Systems **34** (2021)
44. Sun, J., Qu, Q., Wright, J.: Complete dictionary recovery over the sphere i: Overview and the geometric picture. IEEE Transactions on Information Theory **63**(2), 853–884 (2016)
45. Sun, J., Qu, Q., Wright, J.: A geometric analysis of phase retrieval. Foundations of Computational Mathematics **18**(5), 1131–1198 (2018)
46. Sun, R., Luo, Z.Q.: Guaranteed matrix completion via non-convex factorization. IEEE Transactions on Information Theory **62**(11), 6535–6579 (2016)
47. Tong, T., Ma, C., Chi, Y.: Accelerating ill-conditioned low-rank matrix estimation via scaled gradient descent. Journal of Machine Learning Research **22**(150), 1–63 (2021)
48. Tong, T., Ma, C., Chi, Y.: Low-rank matrix recovery with scaled subgradient methods: Fast and robust convergence without the condition number. IEEE Transactions on Signal Processing **69**, 2396–2409 (2021)
49. Tong, T., Ma, C., Prater-Bennette, A., Tripp, E., Chi, Y.: Scaling and scalability: Provable nonconvex low-rank tensor estimation from incomplete measurements. arXiv preprint arXiv:2104.14526 (2021)

50. Tu, S., Boczar, R., Simchowitz, M., Soltanolkotabi, M., Recht, B.: Low-rank solutions of linear matrix equations via procrustes flow. In: International Conference on Machine Learning, pp. 964–973. PMLR (2016)
51. Wainwright, M.J.: High-dimensional statistics: A non-asymptotic viewpoint, vol. 48. Cambridge University Press (2019)
52. Waldspurger, I., Waters, A.: Rank optimality for the burer–monteiro factorization. SIAM journal on Optimization **30**(3), 2577–2602 (2020)
53. Wei, K., Cai, J.F., Chan, T.F., Leung, S.: Guarantees of riemannian optimization for low rank matrix recovery. SIAM Journal on Matrix Analysis and Applications **37**(3), 1198–1222 (2016)
54. Wei, K., Cai, J.F., Chan, T.F., Leung, S.: Guarantees of riemannian optimization for low rank matrix completion. Inverse Problems & Imaging **14**(2) (2020)
55. Yalcin, B., Zhang, H., Lavaei, J., Sojoudi, S.: Factorization approach for low-complexity matrix completion problems: Exponential number of spurious solutions and failure of gradient methods. In: International Conference on Artificial Intelligence and Statistics, pp. 1–9. PMLR (2022)
56. Yi, X., Park, D., Chen, Y., Caramanis, C.: Fast algorithms for robust pca via gradient descent. Advances in neural information processing systems **29** (2016)
57. Zhang, H., Bi, Y., Lavaei, J.: General low-rank matrix optimization: Geometric analysis and sharper bounds. Advances in Neural Information Processing Systems **34** (2021)
58. Zhang, J., Fattahi, S., Zhang, R.: Preconditioned gradient descent for over-parameterized nonconvex matrix factorization. Advances in Neural Information Processing Systems **34** (2021)
59. Zhang, R.Y., Sojoudi, S., Lavaei, J.: Sharp restricted isometry bounds for the inexistence of spurious local minima in nonconvex matrix recovery. J. Mach. Learn. Res. **20**(114), 1–34 (2019)
60. Zheng, Q., Lafferty, J.: A convergent gradient descent algorithm for rank minimization and semidefinite programming from random linear measurements. In: Proceedings of the 28th International Conference on Neural Information Processing Systems-Volume 1, pp. 109–117 (2015)
61. Zhu, Z., Li, Q., Tang, G., Wakin, M.B.: Global optimality in low-rank matrix optimization. IEEE Transactions on Signal Processing **66**(13), 3614–3628 (2018)

A Analysis of the degenerate case

In this section, we provide a detailed analysis on instances with $u^* = 0$. The optimization problem of the instance $\mathcal{MC}(C, 0)$ can be written as

$$(A.1) \qquad \min_{u \in \mathbb{R}^n} \sum_{i,j \in [n]} C_{ij} u_i^2 u_j^2.$$

We prove that problem (A.1) either has multiple global solutions or has no SSCPs.

Theorem A.1 *If $C_{ii} > 0$ for all $i \in [n]$, the instance $\mathcal{MC}(C, 0)$ has no SSCPs. Otherwise if $C_{ii} = 0$ for some $i \in [n]$, the instance $\mathcal{MC}(C, 0)$ has nonzero global solutions.*

Proof We first consider the case when $C_{ii} > 0$ for all $i \in [n]$. Let $u^0 \in \mathbb{R}^n$ be a second-order critical point. By the first-order optimality conditions, it holds that

$$\frac{1}{4} \nabla_i g(u^0; C, 0) = C_{ii}(u_i^0)^3 + \sum_{j \in [n], j \neq i} C_{ij} u_i^0 (u_j^0)^2 = 0, \quad \forall i \in [n].$$

Multiplying u_i^0 on both sides, we have

$$0 = C_{ii}(u_i^0)^4 + \sum_{j \in [n], j \neq i} C_{ij}(u_i^0)^2 (u_j^0)^2 \geq C_{ii}(u_i^0)^4 \geq 0,$$

which implies that $C_{ii}(u_i^0)^4 = 0$. Since $C_{ii} > 0$, it follows that

$$u_i^0 = 0, \quad \forall i \in [n].$$

Hence, $u^0 = 0$ is the unique second-order critical point.

Next, we consider the case when there exists an index i_0 such that $C_{i_0 i_0} = 0$. In this case, define $u^0 \in \mathbb{R}^n$ by

$$u_{i_0}^0 = 1, \quad u_i^0 = 0, \quad \forall i \in [n] \backslash \{i_0\}.$$

Then, we have

$$\left[u^0(u^0)^T\right]_{i_0 i_0} = 1, \quad \left[u^0(u^0)^T\right]_{ij} = 0, \quad \text{otherwise.}$$

Since the (i_0, i_0) entry is not observed, the point u^0 leads to the same measurements as $u^* = 0$. Therefore, u^0 is a nonzero global solution to the instance $\mathcal{MC}(C, 0)$.

$\square$

B Proofs in Section 2

B.1 Proof of Lemma 2.1

Proof Denote the distance between (C, u^*) and $\mathcal{D}$ as

$$\mathbb{T}_\alpha(C, u^*) := \min_{(\tilde{C}, \tilde{u}^*) \in \overline{\mathcal{D}}} \alpha \|C - \tilde{C}\|_1 + (1 - \alpha)\|u^* - \tilde{u}^*\|_1.$$

We fix the pair (C, u^*) and let

$$\eta := \frac{1}{n(n-1)} \sum_{i,j \in [n], i \neq j} C_{ij} \in \left[0, \frac{1}{n(n-1)}\right].$$

Using the condition $\|C\|_1 = 1$, it follows that

$$\theta := \frac{1}{n} \sum_{i \in [n]} C_{ii} = \frac{1}{n}\left(1 - \sum_{i,j \in [n], i \neq j} C_{ij}\right) = \frac{1}{n} - (n-1)\eta \in [0, n^{-1}].$$

Our goal is to prove that

$$\mathbb{T}_\alpha(C, u^*) \leq g(\alpha, \eta).$$

In the remainder of the proof, we upper-bound the distance $\mathbb{T}_\alpha(C, u^*)$ by constructing some instances in $\overline{\mathcal{D}}$.

We first consider those instances in $\overline{\mathcal{D}}$ with a disconnected subgraph $\mathbb{G}_1$. For every $k \in \{2, \ldots, n\}$, let $\mathcal{I}_1$ be a subset of $[n]$ satisfying $|\mathcal{I}_1| = k$ and $\mathcal{I}_0 := [n] \backslash \mathcal{I}_1$. Suppose that $\epsilon > 0$ is a sufficiently small constant. For every $i_0 \in \mathcal{I}_1$, we consider the pair $(\tilde{C}, \tilde{u}^*)$, where

(B.1) $\qquad \tilde{u}_i^* = 0, \quad \forall i \in \mathcal{I}_0; \quad \tilde{u}_i^* = (1 - \epsilon)u_i^* + \epsilon \cdot \frac{\left\|u_{\mathcal{I}_1}^*\right\|_1}{|\mathcal{I}_1|} + \frac{\left\|u_{\mathcal{I}_0}^*\right\|_1}{|\mathcal{I}_1|}, \quad \forall i \in \mathcal{I}_1$

and

$$\tilde{C}_{i_0 j} = \tilde{C}_{j i_0} = 0, \quad \forall j \in \mathcal{I}_1 \backslash \{i_0\};$$
$$\tilde{C}_{ij} = C_{ij} + \frac{2}{n^2 - 2(k-1)} \sum_{j \in \mathcal{I}_1 \backslash \{i_0\}} C_{i_0 j}, \quad \text{otherwise.}$$

By choosing a sufficiently small ϵ, it can be shown that

$$\mathcal{I}_1(\tilde{C}, \tilde{u}^*) = \mathcal{I}_1; \quad \mathcal{I}_0(\tilde{C}, \tilde{u}^*) = \mathcal{I}_0.$$

The node i_0 is disconnected from other nodes in $\mathbb{G}_1(\tilde{C}, \tilde{u}^*)$ and, therefore, $(\tilde{C}, \tilde{u}^*) \in \overline{\mathcal{D}}$. The distance between u^* and $\tilde{u}^*$ is

(B.2) $\qquad \|u^* - \tilde{u}^*\|_1 = 2\left\|u_{\mathcal{I}_0}^*\right\|_1 + 2\epsilon\left\|u_{\mathcal{I}_1}^*\right\|_1 \leq 2\left\|u_{\mathcal{I}_0}^*\right\|_1 + 2\epsilon.$

In addition, the distance between C and $\tilde{C}$ can be calculated as

$$(B.3) \qquad \|C - \tilde{C}\|_1 = 4\sum_{j \in \mathcal{I}_1 \setminus \{i_0\}} C_{i_0 j}.$$

Combining inequalities (B.2) and (B.3), we have

$$(B.4) \qquad \mathbb{T}_\alpha(C, u^*) \le 2(1 - \alpha) \left\| u^*_{\mathcal{I}_0} \right\|_1 + 4\alpha \sum_{j \in \mathcal{I}_1 \setminus \{i_0\}} C_{i_0 j} + 2\epsilon.$$

Taking the average of inequality (B.4) over i_0 for $\mathcal{I}_1$, we have

$$(B.5) \qquad \mathbb{T}_\alpha(C, u^*) \le 2(1 - \alpha) \left\| u^*_{\mathcal{I}_0} \right\|_1 + 4\alpha(k - 1) \sum_{i,j \in \mathcal{I}_1, i \ne j} C_{ij} + 2\epsilon.$$

Then, we take the average of (B.5) over $\mathcal{I}_1$ for all k-element subsets of $[n]$, which leads to

$$\mathbb{T}_\alpha(C, u^*) \le 2(1 - \alpha) \cdot \frac{n - k}{n} + 4\alpha(k - 1)\eta + 2\epsilon.$$

By setting $\epsilon \to 0$, we obtain that

$$(B.6) \qquad \mathbb{T}_\alpha(C, u^*) \le 2(1 - \alpha) \cdot \frac{n - k}{n} + 4\alpha(k - 1)\eta.$$

Since inequality (B.6) is linear in k, the minimum of the right-hand side over $k \in \{2, \ldots, n\}$ is attained by either 2 or n. Hence, it holds that

$$(B.7) \qquad \mathbb{T}_\alpha(C, u^*) \le \min\left\{ 2(1 - \alpha) \cdot \frac{n - 2}{n} + 4\alpha\eta, \, 4\alpha(n - 1)\eta \right\}.$$

Using a similar analysis, we can obtain inequality (B.6) by considering instances in $\overline{\mathcal{D}}$ whose $\mathcal{I}_{00}$ is non-empty.

Finally, we check those instances in $\overline{\mathcal{D}}$ whose $\mathbb{G}_1$ is bipartite. Let $\mathcal{I}_1$ be a subset of $[n]$ satisfying $|\mathcal{I}_1| = 4$, and let $\mathcal{I}_0 = [n] \setminus \mathcal{I}_1$. We define $\tilde{u}^*$ in the same way as (B.1). For every subset $\mathcal{I}_{11} \subset \mathcal{I}_1$ such that $|\mathcal{I}_{11}| = 2$, the new weight matrix is defined as

$$\tilde{C}_{ii} = 0, \quad \forall i \in \mathcal{I}_1; \quad \tilde{C}_{ij} = 0, \quad \forall i, j \in \mathcal{I}_{11}; \quad \tilde{C}_{ij} = 0, \quad \forall i, j \in \mathcal{I}_1 \setminus \mathcal{I}_{11};$$

$$\tilde{C}_{ij} = C_{ij} + \frac{2}{n^2 - 8} \left(\sum_{i \in \mathcal{I}_1} C_{ii} + \sum_{i,j \in \mathcal{I}_{11}, i \ne j} C_{ij} + \sum_{i,j \in \mathcal{I}_1 \setminus \mathcal{I}_{11}, i \ne j} C_{ij} \right).$$

The distance between C and $\tilde{C}$ is

$$\|C - \tilde{C}\|_1 = 2 \left(\sum_{i \in \mathcal{I}_1} C_{ii} + \sum_{i,j \in \mathcal{I}_{11}, i \ne j} C_{ij} + \sum_{i,j \in \mathcal{I}_1 \setminus \mathcal{I}_{11}, i \ne j} C_{ij} \right)$$

Therefore, the maximum distance is bounded by

$$(B.8) \qquad \mathbb{T}_\alpha(C, u^*) \le 2(1 - \alpha) \left\| u^*_{\mathcal{I}_0} \right\|_1$$

$$+ 2\alpha \left(\sum_{i \in \mathcal{I}_1} C_{ii} + \sum_{i,j \in \mathcal{I}_{11}, i \ne j} C_{ij} + \sum_{i,j \in \mathcal{I}_1 \setminus \mathcal{I}_{11}, i \ne j} C_{ij} \right) + 2\epsilon.$$

By taking the average of (B.8) over $\mathcal{I}_{11}$ for all 2-element subsets of $\mathcal{I}_1$, it follows that

$$(B.9) \qquad \mathbb{T}_\alpha(C, u^*) \le 2(1 - \alpha) \left\| u^*_{\mathcal{I}_0} \right\|_1 + 2\alpha \left(\sum_{i \in \mathcal{I}_1} C_{ii} + \frac{1}{3} \sum_{i,j \in \mathcal{I}_1, i \ne j} C_{ij} \right) + 2\epsilon.$$

Furthermore, we take the average of (B.9) over $\mathcal{I}_1$ for all 4-element subsets of $[n]$, which gives

$$\mathbb{T}_\alpha(C, u^*) \leq 2(1 - \alpha) \cdot \frac{k}{n} + 2\alpha\left(4\theta + 4\eta\right) + 2\epsilon.$$

By letting $\epsilon \to 0$, we conclude that

(B.10) $$\mathbb{T}_\alpha(C, u^*) \leq 2(1 - \alpha) \cdot \frac{4}{n} + 2\alpha\left(4\theta + 4\eta\right).$$

By applying a similar technique to subsets of $[n]$ with $1, 2, 3$ elements, the distance can be bounded as

(B.11) $$\mathbb{T}_\alpha(C, u^*) \leq 2(1 - \alpha) \cdot \frac{3}{n} + 2\alpha\left(3\theta + 2\eta\right),$$

$$\mathbb{T}_\alpha(C, u^*) \leq 2(1 - \alpha) \cdot \frac{2}{n} + 2\alpha \cdot 2\theta,$$

$$\mathbb{T}_\alpha(C, u^*) \leq 2(1 - \alpha) \cdot \frac{1}{n} + 2\alpha \cdot \theta.$$

By combining inequalities (B.6), (B.10) and (B.11) and recalling the relation that $\theta = 1/n - (n - 1)\eta$, it follows that

$$\mathbb{T}_\alpha(C, u^*) \leq g(\alpha, \eta).$$

Now, we take the maximum over $C \in \mathbb{S}_{+,1}^{n^2-1}$ and $u^* \in \mathbb{S}_1^{n-1}$, which is equivalent to taking the maximum over $\eta \in \left[0, \frac{1}{n(n-1)}\right]$ in the right-hand side. This yields that

$$\max_{\|C\|_1 = \|u^*\|_1 = 1} \mathbb{T}_\alpha(C, u^*) \leq \max_{c \in \left[0, \frac{1}{n(n-1)}\right]} g(\alpha, c).$$

This completes the proof. $\square$

B.2 Proof of Lemma 2.2

Proof Let $\eta \in \left[0, \frac{1}{n(n-1)}\right]$ and define the pair (C, u^*) according to

$$u_i^* := \frac{1}{n}, \ C_{ii} := \frac{1}{n} - (n - 1)\eta, \quad \forall i \in [n]; \quad C_{ij} := \eta, \quad \forall i, j \in [n] \quad \text{s.t.} \quad i \neq j.$$

Our goal is to prove that

$$\mathbb{T}_\alpha(C, u^*) \geq g(\alpha, \eta).$$

Suppose that $(\tilde{C}, \tilde{u}^*) \in \overline{\mathcal{D}}$ attains the distance $\mathbb{T}_\alpha(C, u^*)$, namely,

$$\mathbb{T}_\alpha(C, u^*) = \alpha\|C - \tilde{C}\|_1 + (1 - \alpha)\|u^* - \tilde{u}^*\|_1.$$

We analyze three different cases.

Case I. We first consider the case when $\mathbb{G}_1(\tilde{C}, \tilde{u}^*)$ is disconnected. Denote $k := |\mathcal{I}_1(\tilde{C}, \tilde{u}^*)|$. The distance between u^* and $\tilde{u}^*$ is lower-bounded by

(B.12) $$\|u^* - \tilde{u}^*\|_1 \geq 2\|u_{\mathcal{I}_0(\tilde{C},\tilde{u}^*)}^* - \tilde{u}_{\mathcal{I}_0(\tilde{C},\tilde{u}^*)}^*\|_1 = 2\|u_{\mathcal{I}_0(\tilde{C},\tilde{u}^*)}^*\|_1 = \frac{2(n - k)}{n}.$$

Since there are k nodes in $\mathbb{G}_1(\tilde{C}, \tilde{u}^*)$, we need to eliminate at least $k - 1$ edges that are not self-loops to make the graph disconnected. Therefore, at least $2(k - 1)$ non-diagonal weights of $\tilde{C}$ are 0 and the distance between C and $\tilde{C}$ is at least

(B.13) $$\|C - \tilde{C}\|_1 \geq 2 \cdot 2(k - 1)\eta = 4(k - 1)\eta.$$

Combining inequalities (B.12) and (B.13), we obtain that

(B.14) $$\mathbb{T}_\alpha(C, u^*) \geq 2(1 - \alpha) \cdot \frac{n - k}{n} + 4\alpha(k - 1)\eta.$$

Case II. For the case when $\mathcal{I}_{00}(\tilde{C}, \tilde{u}^*)$ is not empty, similar estimations as *Case I* can be derived and inequality (B.14) also holds true.

Case III. Finally, we consider the case when $\mathbb{G}_1(\tilde{C}, \tilde{u}^*)$ is bipartite. Denote $k := |\mathcal{I}_1(\tilde{C}, \tilde{u}^*)|$. If $k \geq 5$, we need to eliminate at least $k - 1$ edges that are not self-loops to make the graph bipartite. Thus, we can follow the same proof as *Case I* to arrive at inequality (B.14). If $k = 4$, we need to eliminate at least 2 edges that are not self-loops and 4 self-loops to make the graph bipartite. Therefore, at least 4 non-diagonal weights and 4 diagonal weights of $\tilde{C}$ are 0, and the distance between C and $\tilde{C}$ is at least

$$(\text{B.15}) \qquad \|C - \tilde{C}\|_1 \geq 2\left[4\eta + 4\left(\frac{1}{n} - (n-1)\eta\right)\right] = 2\left[\frac{4}{n} - (4n-8)\eta\right].$$

Combining inequalities (B.12) and (B.15) yields that

$$(\text{B.16}) \qquad \mathbb{T}_\alpha(C, u^*) \geq 2(1-\alpha) \cdot \frac{n-4}{n} + 2\alpha\left[\frac{4}{n} - (4n-8)\eta\right].$$

The cases when $k = 1, 2, 3$ can be analyzed similarly, leading to

$$(\text{B.17}) \qquad \mathbb{T}_\alpha(C, u^*) \geq 2(1-\alpha) \cdot \frac{n-3}{n} + 2\alpha\left[\frac{3}{n} - (3n-5)\eta\right],$$

$$\mathbb{T}_\alpha(C, u^*) \geq 2(1-\alpha) \cdot \frac{n-2}{n} + 2\alpha\left[\frac{2}{n} - (2n-2)\eta\right],$$

$$\mathbb{T}_\alpha(C, u^*) \geq 2(1-\alpha) \cdot \frac{n-1}{n} + 2\alpha\left[\frac{1}{n} - (n-1)\eta\right].$$

By combining *Cases I-III*, it follows that

$$\mathbb{T}_\alpha(C, u^*) \geq g(\alpha, \eta).$$

Choosing η to be the maximizer

$$\eta^* := \underset{c \in \left[0, \frac{1}{n(n-1)}\right]}{\arg\max} \; g(\alpha, c),$$

we have

$$\mathbb{T}_\alpha(C, u^*) \geq \underset{c \in \left[0, \frac{1}{n(n-1)}\right]}{\max} g(\alpha, c).$$

Taking the maximum over $C \in \mathbb{S}_{+,1}^{n^2-1}$ and $u^* \in \mathbb{S}_1^{n-1}$ gives rise to the desired conclusion. $\qquad\square$

C Proofs in Section 3

C.1 Proof of Lemma 3.1

Proof We first prove the necessity part. Suppose that the instance $\mathcal{MC}(C, u^*)$ satisfies the δ-RIP$_{2,2}$ condition with the constants c_1 and c_2. For every $i, j \in [n]$, we define the matrix $E^{ij} \in \mathbb{R}^{n \times n}$ of rank at most 2 as

$$(\text{C.1}) \qquad E_{ij}^{ij} = E_{ji}^{ij} = 1, \quad E_{k\ell}^{ij} = 0, \quad \text{otherwise}.$$

Using the facts that $\nabla^2 f(M; M^*) = C$ and C is symmetric, it holds that

$$\left[\nabla^2 f(M; M^*)\right](E^{ij}, E^{ij}) = C_{ij} + C_{ji} = 2C_{ij}, \quad \|E^{ij}\|_F^2 = 2.$$

By the RIP$_{2,2}$ condition (3.1), we have

$$2c_1 \leq 2C_{ij} \leq 2c_2, \quad \forall i,j \in [n].$$

which leads to

$$\frac{\min_{i,j \in [n]} C_{ij}}{\max_{i,j \in [n]} C_{ij}} \geq \frac{c_1}{c_2} = \frac{1-\delta}{1+\delta}.$$

Now, we consider the sufficiency part. We aim to prove that

(C.2) $$\min_{i,j \in [n]} C_{ij} \cdot \|K\|_F^2 \leq \left[\nabla^2 f(M; M^*)\right](K, K) \leq \max_{i,j \in [n]} C_{ij} \cdot \|K\|_F^2.$$

Recalling that $\nabla^2 f(M; M^*) = C$, we can calculate that

$$\left[\nabla^2 f(M; M^*)\right](K, K) = \sum_{i,j \in [n]} C_{ij} K_{ij}^2.$$

The inequality (C.2) directly follows from the above inequality and the condition that $C_{ij} \geq 0$ for all $i,j \in [n]$. Therefore, we can choose

$$c_1 := \min_{i,j \in [n]} C_{ij}, \quad c_2 := \max_{i,j \in [n]} C_{ij}.$$

Then, the δ-RIP$_{2,2}$ condition holds under the assumption of the theorem. $\qquad\square$

C.2 Proof of Lemma 3.2

Proof Assume without loss of generality that

$$|u_i^*| > 0, \quad \forall i \in [\ell]; \quad u_i^* = 0, \quad \forall i \in \{\ell+1, \ldots, n\}.$$

By the definition (1.6), we have

$$(u_i^*)^2 \leq \frac{\mu}{n} \|u^*\|_2^2 = \frac{\mu}{n} \sum_{i \in [\ell]} (u_i^*)^2, \quad \forall i \in [\ell].$$

Summing over $i \in [\ell]$, we obtain that

$$\sum_{i \in [\ell]} (u_i^*)^2 \leq \frac{\ell \mu}{n} \sum_{i \in [\ell]} (u_i^*)^2,$$

which implies that $\ell \geq n/\mu$. Let

$$c_i := |u_i^*|/\|u^*\|_2, \quad \forall i \in [\ell].$$

The μ-incoherence condition implies that

(C.3) $$c_i \in (0, \sqrt{\mu/n}], \quad \forall i \in [\ell].$$

In addition, it holds that

$$\|u^*\|_2^2 = \sum_{i \in [\ell]} (u_i^*)^2 = \sum_{i \in [\ell]} c_i^2 \|u^*\|_2^2,$$

$$1 = \|u^*\|_1 = \sum_{i \in [\ell]} |u_i^*| = \sum_{i \in [\ell]} c_i \|u^*\|_2,$$

which implies that

$$\sum_{i \in [\ell]} c_i^2 = 1, \quad \sum_{i \in [\ell]} c_i = \|u^*\|_2^{-1}.$$

Combined with (C.3), it follows that

$$\|u^*\|_2^{-1} = \sum_{i \in [\ell]} c_i \geq \sqrt{\frac{n}{\mu}} \cdot \sum_{i \in [\ell]} c_i^2 = \sqrt{\frac{n}{\mu}}.$$

Therefore,

$$|u_i^*| = c_i \|u^*\|_2 \leq \sqrt{\mu/n} \cdot \sqrt{\mu/n} = \mu/n.$$

$\qquad\square$

C.3 Proof of Theorem 3.8

Proof We pursue a similar way as in Theorem 3.7 to construct spurious solutions. By Lemma 3.7, we only need to show that problem (3.12) has at least $\binom{m}{(m-1)/2}$ SSCPs whose Hessian matrices are positive definite and whose components are nonzero. Let $k := (m-1)/2 \in \mathbb{Z}$. We first choose a subset

$$\mathcal{I} \subset [m], \quad |\mathcal{I}| = k.$$

Then, we consider the point $x \in \mathbb{R}^m$, where

$$u_i = y_1, \quad \forall i \in \mathcal{I}, \quad u_i = y_2, \quad \forall i \notin \mathcal{I},$$

where y_1 and y_2 are real numbers such that

(C.4)
$$(1+k\epsilon)(1+2k\epsilon)[(1-\epsilon)y_2^2]^3 - 2(1+k\epsilon)(1+(k-1)\epsilon)[(1-\epsilon)y_2^2]^2$$
$$+ (1+k\epsilon)(1+(k-1)\epsilon)(2k^2\epsilon^2 + 2k\epsilon^2 - k\epsilon - \epsilon + 1)[(1-\epsilon)y_2^2]$$
$$- k^2\epsilon^2(1+(k-1)\epsilon)(1-\epsilon)^2 = 0,$$
$$y_1 = \frac{y_2}{k\epsilon} \cdot \frac{(1+k\epsilon)[(1-\epsilon)y_2^2] - (k^2\epsilon^2 + (k-1)\epsilon + 1)}{[(1-\epsilon)y_2^2] + (1+(k-1)\epsilon)}.$$

We first assume the existence of the constants y_1 and y_2. After some direct calculations, one can show that the conditions in (C.4) imply the first-order optimality condition of the instance $\mathcal{MC}(C^\epsilon, u^*)$, i.e.,

$$y_1^3 - y_1 + \epsilon[(k-1)y_1^2 + (k+1)y_2^2]y_1 - \epsilon[(k-1)y_1 + (k+1)y_2] = 0,$$
$$y_2^3 - y_2 + \epsilon[ky_1^2 + ky_2^2]y_2 - \epsilon[ky_1 + ky_2] = 0.$$

Therefore, the point x is a first-order critical point of the instance $\mathcal{MC}(C^\epsilon, u^*)$. In addition, the following relations result from the condition (C.4):

(C.5)
$$(1-\epsilon)y_1y_2(y_1+y_2) = -\epsilon[ky_1 + (k+1)y_2],$$
$$(1-\epsilon)(y_1^2 + y_1y_2 + y_2^2 - 1) = -\epsilon[ky_1^2 + (k+1)y_2^2].$$

Now, we prove the existence of y_1, y_2 and estimate their values. We note that the first equation in (C.4) is a third-order polynomial equation for $(1-\epsilon)y_2^2$, which has at least one real root. To show that the equation has a positive root, we observe that the coefficient of the third-order term is $(1+k\epsilon)(1+2k\epsilon) > 0$ and the value at zero is $-k^2\epsilon^2(1+(k-1)\epsilon)(1-\epsilon)^2 < 0$. Therefore, the polynomial equation for $(1-\epsilon)y_2^2$ has at least one positive root and y_2 is well defined. We provide a more accurate estimate to y_1 and y_2, namely, we show that there exists a solution (y_1, y_2) to equations (C.4) such that

$$y_1 \in [-2, -3/5], \quad y_2 \in [1/2, 1].$$

Define the polynomial function

$$g(z) := (1+k\epsilon)(1+2k\epsilon)z^3 - 2(1+k\epsilon)(1+(k-1)\epsilon)z^2$$
$$+ (1+k\epsilon)(1+(k-1)\epsilon)(2k^2\epsilon^2 + 2k\epsilon^2 - k\epsilon - \epsilon + 1)z - k^2\epsilon^2(1+(k-1)\epsilon)(1-\epsilon)^2.$$

We first estimate $g(1-(2k+1)\epsilon)$ as follows:

$$g(1-(2k+1)\epsilon)$$
$$= (1+k\epsilon)[1-(2k+1)\epsilon]\Big[(1+2k\epsilon)[1-(2k+1)\epsilon]^2 - 2[1+(k-1)\epsilon][1-(2k+1)\epsilon]$$
$$+ [1+(k-1)\epsilon][1-(k+1)\epsilon + 2k(k+1)\epsilon^2]\Big] - k^2\epsilon^2(1+(k-1)\epsilon)(1-\epsilon)^2$$
$$= (1+k\epsilon)[1-(2k+1)\epsilon]\Big[k^2\epsilon^2 + 2k^2(5k+4)\epsilon^3\Big] - k^2\epsilon^2(1+(k-1)\epsilon)(1-\epsilon)^2$$
$$\geq k^2\epsilon^2(1+k\epsilon)[1-(2k+1)\epsilon][1+2(5k+4)\epsilon] - k^2\epsilon^2(1+k\epsilon)(1-\epsilon)^2$$
$$= k^2\epsilon^2(1+k\epsilon)[(8k+9)\epsilon - [2(2k+1)(5k+4)+1]\epsilon^2]$$
$$\geq k^2\epsilon^2(1+k\epsilon)[(8k+8)\epsilon - 20(k+1)^2\epsilon^2] > 0,$$

where the last inequality is due to $(k+1)\epsilon = (n+1)\epsilon/2 < 2/5$. Next, we estimate $g(1 - (3k/2 + 1)\epsilon)$ as follows:

$$
\begin{aligned}
&g(1 - (3k/2 + 1)\epsilon) \\
&= (1 + k\epsilon)[1 - (k+1)\epsilon]\Big[(1 + 2k\epsilon)[1 - (3k/2 + 1)\epsilon]^2 - 2[1 + (k-1)\epsilon][1 - (3k/2 + 1)\epsilon] \\
&\quad + [1 + (k-1)\epsilon][1 - (k+1)\epsilon + 2k(k+1)\epsilon^2]\Big] - k^2\epsilon^2(1 + (k-1)\epsilon)(1 - \epsilon)^2 \\
&= (1 + k\epsilon)[1 - (3k/2 + 1)\epsilon]\Big[k^2\epsilon^2/4 + k^2(13k/2 + 6)\epsilon^3\Big] - k^2\epsilon^2(1 + (k-1)\epsilon)(1 - \epsilon)^2 \\
&= k^2\epsilon^2(1 + k\epsilon)[1 - (3k/2 + 1)\epsilon][1/4 + (13k/2 + 6)\epsilon] - k^2\epsilon^2(1 + (k-1)\epsilon)(1 - \epsilon)^2 \\
&\leq k^2\epsilon^2(1 + k\epsilon)[1 - (3k/2 + 1)\epsilon][1/4 + (13k/2 + 6)\epsilon] - k^2\epsilon^2 \cdot [(1 + k\epsilon)/2] \cdot (1 - \epsilon)^2 \\
&\leq k^2\epsilon^2(1 + k\epsilon)\Big[[1 - (3k/2 + 1)\epsilon][1/4 + (13k/2 + 6)\epsilon] - (1 - \epsilon)^2/2\Big] \\
&= k^2\epsilon^2(1 + k\epsilon)\Big[-1/4 + (49k/8 + 27/4)\epsilon - (39k^2/4 + 31k/2 + 13/2)\epsilon^2\Big] \\
&\leq k^2\epsilon^2(1 + k\epsilon)\Big[-1/4 + 27(k+1)/4\epsilon - 39(k+1)^2\epsilon^2/4\Big] < 0,
\end{aligned}
$$

where the last inequality is in light of $(k+1)\epsilon = (n+1)\epsilon/2 < 1/26$. Combining the above two estimates, we conclude that there exists a solution y_2 to the first equation in (C.4) such that

$$
\text{(C.6)} \qquad (1 - \epsilon)y_2^2 \in [1 - (2k+1)\epsilon, 1 - (3k/2 + 1)\epsilon].
$$

Hence,

$$
\text{(C.7)} \qquad y_2 \leq \sqrt{\frac{1 - (3k/2 + 1)\epsilon}{1 - \epsilon}} \leq 1
$$

and

$$
\text{(C.8)} \qquad y_2 \geq \sqrt{\frac{1 - (2k+1)\epsilon}{1 - \epsilon}} \geq \sqrt{1 - (2k+1)\epsilon} \geq \frac{1}{2}.
$$

Now, we use the second equation in (C.4) to estimate y_1, which leads to

$$
\begin{aligned}
&\frac{(1 + k\epsilon)[(1 - \epsilon)y_2^2] - (k^2\epsilon^2 + (k-1)\epsilon + 1)}{k\epsilon} \\
&\geq \frac{(1 + k\epsilon)[1 - (2k+1)\epsilon] - (k^2\epsilon^2 + (k-1)\epsilon + 1)}{k\epsilon} \\
&= -2 - (3k+1)\epsilon
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{(1 + k\epsilon)[(1 - \epsilon)y_2^2] - (k^2\epsilon^2 + (k-1)\epsilon + 1)}{k\epsilon} \\
&\leq \frac{(1 + k\epsilon)[1 - (3k/2 + 1)\epsilon] - (k^2\epsilon^2 + (k-1)\epsilon + 1)}{k\epsilon} \\
&= -\frac{3}{2} - \left(\frac{5k}{2} + 1\right)\epsilon.
\end{aligned}
$$

On the other hand, we have

$$
\frac{y_2}{[(1 - \epsilon)y_2^2] + (1 + (k-1)\epsilon)} = \frac{1}{(1 - \epsilon)(y_2 + y_2^{-1}) + k\epsilon} \leq \frac{1}{2(1 - \epsilon) + k\epsilon}.
$$

Using the bound in (C.6), it holds that

$$y_2 \geq \sqrt{\frac{1 - (2k+1)\epsilon}{1-\epsilon}} \geq \frac{1 - (2k+1)\epsilon}{1-\epsilon} = \frac{1}{2} - \frac{1 - (4k+1)\epsilon}{2(1-\epsilon)} \geq \frac{1}{2}.$$

Therefore,

$$\frac{y_2}{[(1-\epsilon)y_2^2] + (1 + (k-1)\epsilon)} = \frac{1}{(1-\epsilon)(y_2 + y_2^{-1}) + k\epsilon} \geq \frac{1}{2.5(1-\epsilon) + k\epsilon}.$$

Combining the above inequalities and the second equation in (C.4) yields that

$$(C.9) \qquad y_1 \geq \frac{-2 - (3k+1)\epsilon}{2(1-\epsilon) + k\epsilon} \geq -\left(1 + \frac{5\epsilon}{1-\epsilon}\right) \geq -2$$

and

$$(C.10) \qquad y_1 \leq \frac{-3/2 - (5k/2+1)\epsilon}{2.5(1-\epsilon) + k\epsilon} \leq -\frac{1.5}{2.5} = -\frac{3}{5},$$

where the last inequality in (C.9) results from $\epsilon \leq 1/(3(k+1)) \leq 1/6$. In summary, inequalities (C.7)-(C.10) lead to

$$y_1 \in [-2, -3/5], \quad y_2 \in [1/2, 1].$$

We then prove that $y_1 + 2y_2 > y_2 \geq 0.5$, which is equivalent to

$$\frac{y_2}{k\epsilon} \cdot \frac{(1+k\epsilon)[(1-\epsilon)y_2^2] - (k^2\epsilon^2 + (k-1)\epsilon + 1)}{[(1-\epsilon)y_2^2] + (1 + (k-1)\epsilon)} + y_2 > 0.$$

Since $y_2 > 0$, we only need to prove that

$$0 < (1+k\epsilon)[(1-\epsilon)y_2^2] - (k^2\epsilon^2 + (k-1)\epsilon + 1) + k\epsilon \Big[[(1-\epsilon)y_2^2] + (1 + (k-1)\epsilon) \Big]$$

$$= (1+2k\epsilon)[(1-\epsilon)y_2^2] - (k^2\epsilon^2 + (k-1)\epsilon + 1) + k\epsilon(1 + (k-1)\epsilon).$$

Using inequality (C.6), it suffices to show that

$$(1+2k\epsilon)[1 - (3k/2+1)\epsilon)] + k\epsilon(1 + (k-1)\epsilon) > 1 + (k-1)\epsilon + k^2\epsilon^2$$

$$\iff \frac{1}{2}k\epsilon > 3k\left(k + \frac{3}{2}\right)\epsilon^2 \iff 3(2k+3)\epsilon < 1 \iff 6(k+1)\epsilon < 1,$$

where the last inequality holds since $(k+1)\epsilon = (n+1)\epsilon/2 < 1/6$.

Now, we verify the second-order sufficient optimality condition. For every $c \in \mathbb{R}^m \setminus \{0\}$, we calculate that

$$c^T H(x; \epsilon)c = \sum_{i \in \mathcal{I}} \left[3y_1^2 - 1 + \epsilon((k-1)y_1^2 + (k+1)y_2^2)\right] c_i^2$$

$$+ \sum_{i \notin \mathcal{I}} \left[3y_2^2 - 1 + \epsilon(ky_1^2 + ky_2^2)\right] c_i^2 + \sum_{i,j \in \mathcal{I}, i \neq j} \epsilon \left(2y_1^2 - 1\right) c_i c_j$$

$$+ \sum_{i,j \notin \mathcal{I}, i \neq j} \epsilon \left(2y_2^2 - 1\right) c_i c_j + 2 \sum_{i \in \mathcal{I}, j \notin \mathcal{I}} \epsilon \left(2y_1 y_2 - 1\right) c_i c_j$$

$$= \left[3y_1^2 - 1 + \epsilon((k-1)y_1^2 + (k+1)y_2^2) - \epsilon \left(2y_1^2 - 1\right)\right] \sum_{i \in \mathcal{I}} c_i^2$$

$$+ \left[3y_2^2 - 1 + \epsilon(ky_1^2 + ky_2^2) - \left(2y_2^2 - 1\right)\right] \sum_{i \notin \mathcal{I}} c_i^2$$

$$+ \epsilon \left(2y_1^2 - 1\right) \left(\sum_{i \in \mathcal{I}} c_i\right)^2 + \epsilon \left(2y_2^2 - 1\right) \left(\sum_{i \notin \mathcal{I}} c_i\right)^2$$

$$+ 2\epsilon \left(2y_1 y_2 - 1\right) \left(\sum_{i \in \mathcal{I}} c_i\right) \left(\sum_{i \notin \mathcal{I}} c_i\right).$$

Using the Cauchy inequality, the above expression is positive if and only if

$$\left[3y_1^2 - 1 + \epsilon((k-1)y_1^2 + (k+1)y_2^2) - \epsilon\left(2y_1^2 - 1\right)\right] \cdot \frac{1}{k}\left(\sum_{i \in \mathcal{I}} c_i\right)^2$$

$$+ \left[3y_2^2 - 1 + \epsilon(ky_1^2 + ky_2^2) - \left(2y_2^2 - 1\right)\right] \cdot \frac{1}{k+1}\left(\sum_{i \notin \mathcal{I}} c_i\right)^2$$

$$+ \epsilon\left(2y_1^2 - 1\right)\left(\sum_{i \in \mathcal{I}} c_i\right)^2 + \epsilon\left(2y_2^2 - 1\right)\left(\sum_{i \notin \mathcal{I}} c_i\right)^2$$

$$+ 2\epsilon\left(2y_1 y_2 - 1\right)\left(\sum_{i \in \mathcal{I}} c_i\right)\left(\sum_{i \notin \mathcal{I}} c_i\right) > 0.$$

We denote

$$A := \sum_{i \in \mathcal{I}} c_i, \quad B := \sum_{i \notin \mathcal{I}} c_i.$$

Then, the second-order sufficient condition is equivalent to

$$\left[3y_1^2 - 1 + \epsilon((k-1)y_1^2 + (k+1)y_2^2) - \epsilon\left(2y_1^2 - 1\right)\right] \cdot \frac{1}{k}A^2$$

$$+ \left[3y_2^2 - 1 + \epsilon(ky_1^2 + ky_2^2) - \left(2y_2^2 - 1\right)\right]$$

$$\cdot \frac{1}{k+1}B^2 + 2\epsilon(y_1 A + y_2 B)^2 - \epsilon(A+B)^2 > 0.$$

The above inequality is a quadratic inequality in A and B, which can be rewritten as

$$\left[\frac{1}{k}\left[3y_1^2 - 1 + \epsilon((k-1)y_1^2 + (k+1)y_2^2) - \epsilon\left(2y_1^2 - 1\right)\right] + \epsilon(2y_1^2 - 1)\right]A^2$$

$$+ 2\epsilon(2y_1 y_2 - 1)AB$$

$$+ \left[\frac{1}{k+1}\left[3y_2^2 - 1 + \epsilon(ky_1^2 + ky_2^2) - \epsilon\left(2y_2^2 - 1\right)\right] + \epsilon(2y_2^2 - 1)\right]B^2 > 0.$$

Therefore, the positivity condition can be verified through the discriminant, namely,

$$\epsilon^2(2y_1 y_2 - 1)^2 < \left[\frac{1}{k}\left[3y_1^2 - 1 + \epsilon((k-1)y_1^2 + (k+1)y_2^2) - \epsilon\left(2y_1^2 - 1\right)\right] + \epsilon(2y_1^2 - 1)\right]$$

$$\cdot \left[\frac{1}{k+1}\left[3y_2^2 - 1 + \epsilon(ky_1^2 + ky_2^2) - \epsilon\left(2y_2^2 - 1\right)\right] + \epsilon(2y_2^2 - 1)\right].$$

Using the second property in (C.5), the above condition can be simplified into

$$-(1-\epsilon)^2(y_2 - y_1)^2(2y_1 + y_2)(y_1 + 2y_2) + (k+1)\epsilon(1-\epsilon)(2y_2^2 - 1)(y_1 - y_2)(2y_1 + y_2)$$

$$+ k\epsilon(1-\epsilon)(2y_1^2 - 1)(y_2 - y_1)(y_1 + 2y_2) > k(k+1)\epsilon^2(y_1 - y_2)^2.$$

Since $y_2 > y_1$, it suffices to have

$$-(1-\epsilon)^2(y_2 - y_1)(2y_1 + y_2)(y_1 + 2y_2) - (k+1)\epsilon(1-\epsilon)(2y_2^2 - 1)(2y_1 + y_2)$$

$$+ k\epsilon(1-\epsilon)(2y_1^2 - 1)(y_1 + 2y_2) > k(k+1)\epsilon^2(y_2 - y_1).$$

We can estimate that

$$-(1-\epsilon)^2(y_2 - y_1)(2y_1 + y_2)(y_1 + 2y_2) - (k+1)\epsilon(1-\epsilon)(2y_2^2 - 1)(2y_1 + y_2)$$

$$+ k\epsilon(1-\epsilon)(2y_1^2 - 1)(y_1 + 2y_2) - k(k+1)\epsilon^2(y_2 - y_1)$$

$$\geq [1 - (k+1)\epsilon]^2 \cdot 1.1 \cdot 0.2 \cdot 0.5 - (k+1)\epsilon \cdot 1 \cdot 0.5 \cdot 2 - (k+1)\epsilon \cdot 1 \cdot 0.64 \cdot 1.4 - (k+1)^2\epsilon^2 \cdot 3$$

$$= 0.11[1 - (k+1)\epsilon]^2 - 1.896(k+1)\epsilon - 3(k+1)^2\epsilon^2 = 0.11 - 2.116(k+1)\epsilon - 2.89(k+1)^2\epsilon^2$$

$$\geq 0.11 - 2.116(k+1)\epsilon - 2.89(k+1)^2\epsilon^2 > 0,$$

where the last inequality is due to $(k+1)\epsilon = (n+1)\epsilon/2 < 1/26$. Thus, we have shown that the Hessian matrix is positive definite and the point x is a SSCP.

To count the number of spurious solutions, we notice that the subset $\mathcal{I}$ has $\binom{m}{(m+1)/2}$ different choices. Hence, the total number of SSCPs is at least $\binom{m}{(m+1)/2}$. The estimate on the combinatorial number follows from $\binom{n}{k} \geq (n/k)^k$. $\qquad\square$

C.4 Proof of Lemma 3.8

Proof We fix ϵ, α and m in the proof. Let $\mathcal{MC}(\tilde{C}, \tilde{u}^*)$ be an instance that attains the minimum in (2.2) and $\ell := |\mathcal{I}_1(\tilde{C}, \tilde{u}^*)|$. Denote

$$d_\alpha := \alpha\|C - \tilde{C}\|_1 + (1-\alpha)\|u^* - \tilde{u}^*\|_1.$$

Then, we investigate three different cases.

Case I. Suppose that $\mathbb{G}_1(\tilde{C}, \tilde{u}^*)$ is disconnected. In this case, at least $2(\ell - 1)$ non-diagonal entries of $\tilde{C}$ are equal to 0. This implies that

(C.11) $$\|C^\epsilon - \tilde{C}\|_1 \geq 4(\ell - 1) \cdot (\epsilon/Z_\epsilon).$$

Case II. The case when $\mathcal{I}_{00}(\tilde{C}, \tilde{u}^*)$ is non-empty can be analyzed similarly as *Case I* and the inequality (C.11) holds. We omit the proof for brevity.

Case III. Finally, we consider the case when $\mathbb{G}_1(\tilde{C}, \tilde{u}^*)$ is bipartite. If $\ell \geq 5$, at least $2(\ell - 1)$ non-diagonal entries of $\tilde{C}$ are equal to 0 and inequality (C.11) holds. If $\ell = 4$, at least 4 non-diagonal entries and 4 diagonal entries of $\tilde{C}$ are equal to 0. Hence, we have

(C.12) $$\|C^\epsilon - \tilde{C}\|_1 \geq 8 \cdot \frac{\epsilon}{Z_\epsilon} + 8 \cdot \frac{1}{Z_\epsilon} = \frac{8\epsilon + 8}{Z_\epsilon} \geq \frac{12\epsilon}{Z_\epsilon}.$$

Similarly, it follows from analyzing the cases with $\ell = 1, 2, 3$ that

(C.13) $$\|C^\epsilon - \tilde{C}\|_1 \geq (4\epsilon + 6)/(Z_\epsilon) \geq 8\epsilon/Z_\epsilon,$$
$$\|C^\epsilon - \tilde{C}\|_1 \geq 4/Z_\epsilon \geq 4\epsilon/Z_\epsilon,$$
$$\|C^\epsilon - \tilde{C}\|_1 \geq 2/Z_\epsilon.$$

Combining inequalities (C.11), (C.12) and (C.13), we know that

(C.14) $$\|C^\epsilon - \tilde{C}\|_1 \geq N(\ell)/Z_\epsilon,$$

where $N(\ell) := 4(\ell - 1)\epsilon$ if $\ell \geq 2$ and $N(1) := 2$.

Now, we consider the optimal choice of $\tilde{u}^*$. Since the distance in (C.14) is increasing in ℓ, it is not optimal to choose $\ell > m$. For every $\ell \in [m]$, at least $m - \ell$ of the first m entries of $\tilde{u}^*$ are 0. Hence, we have the lower bound

(C.15) $$\|u^* - \tilde{u}^*\|_1 \geq 2(m - \ell) \cdot m^{-1}.$$

Combining inequalities (C.14) and (C.15), we have

$$d_\alpha \geq \frac{N(\ell) \cdot \alpha}{Z_\epsilon} + \frac{2(1-\alpha)(m-\ell)}{m}.$$

Taking the minimum over $\ell \in [m]$ leads to

$$d_\alpha \geq \min_{\ell \in [m]} \left[\frac{N(\ell) \cdot \alpha}{Z_\epsilon} + \frac{2(1-\alpha)(m-\ell)}{m} \right].$$

We note that the above inequality indeed attains equality with a suitable choice of $\tilde{C}$ and $\tilde{u}^*$. For all $\ell \geq 2$, we can set $\tilde{u}_i^* = 0$ for all $i \in \{\ell+1, m\}$ and make node 1 disconnected from nodes $\{2, \ldots, \ell\}$. If $\ell = 1$, we can remove the self-loop at node 1. Therefore, it holds that

$$d_\alpha = \min_{\ell \in [m]} \left[\frac{\alpha N(\ell)}{Z_\epsilon} + \frac{2(1-\alpha)(m-\ell)}{m} \right].$$

The minimum in the above equality is attained at one of the points $1, 2, m$, which gives

$$d_\alpha = \min \left\{ \frac{2\alpha}{Z_\epsilon} + \frac{2(1-\alpha)(m-1)}{m}, \frac{4\alpha\epsilon}{Z_\epsilon} + \frac{2(1-\alpha)(m-2)}{m}, \frac{4\alpha(m-1)\epsilon}{Z_\epsilon} \right\}.$$

Since each component in the minimization is an increasing function in ϵ, the distance d_α is also increasing in ϵ. Results for $\mathbb{D}_\alpha(C^\epsilon, u^*)$ follow accordingly by taking the inverse of d_α.

Since the closed form expression of $\mathbb{D}_\alpha(C^\epsilon, u^*)$ is the minimum of three monotone functions in ϵ, the complexity metric is the negative of a unimodal function. For every $\epsilon \leq 1/(2m)$, we can prove that

$$\frac{2\alpha}{Z_\epsilon} + \frac{2(1-\alpha)(m-1)}{m} > \min \left\{ \frac{4\alpha\epsilon}{Z_\epsilon} + \frac{2(1-\alpha)(m-2)}{m}, \frac{4\alpha(m-1)\epsilon}{Z_\epsilon} \right\}.$$

Therefore, in the regime $[0, 1/2]$, the complexity metric $\mathbb{D}_\alpha(C^\epsilon, u^*)$ is the minimum of two strictly decreasing functions and, thus, is also strictly decreasing in ϵ. □

D Proofs in Section 4

D.1 Proof of Lemma 4.1

Proof Without loss of generality, we assume that

$$u_i^0 = 1/n, \quad \forall i \in [n].$$

We first consider the scaled problem instance

$$(D.1) \qquad \min_{x \in \mathbb{R}^n} \sum_{i,j \in [n], i \neq j} (x_i x_j - 1)^2.$$

We denote the gradient and the Hessian matrix of problem (D.1) as $g(x) \in \mathbb{R}^n$ and $H(x) \in \mathbb{R}^{n \times n}$, respectively. Then, we can calculate that

$$\frac{1}{4} g_i(x) = -x_i^3 + (\|x\|_2^2 + 1)x_i - \sum_{k \in [n]} x_k, \quad \forall i \in [n];$$
$$\frac{1}{4} H_{ii}(x) = \sum_{k \in [n], k \neq i} x_k^2, \quad \frac{1}{4} H_{ij}(x) = 2x_i x_j - 1, \quad \forall i, j \in [n].$$

Let c be a small positive constant and define $\epsilon := c/n$. Suppose that $x \in \mathbb{R}^n$ satisfies

$$(D.2) \qquad \|g(x)\|_\infty < 4\epsilon.$$

Then, we study three different cases.

Case I. We first consider the case when $\sum_{i \in [n]} x_i > 2\epsilon$. For all $i \in [n]$, the condition (D.2) implies that

$$(D.3) \qquad \frac{1}{4}|g_i(x)| = \left| \left(\sum_{j \in [n], j \neq i} x_j^2 + 1 \right) x_i - \sum_{j \in [n]} x_j \right| < \epsilon.$$

If $x_i \leq \epsilon$, it holds that

$$\left(\sum_{j \in [n], j \neq i} x_j^2 + 1 \right) x_i - \sum_{j \in [n]} x_j \leq x_i - \sum_{j \in [n]} x_j < -\epsilon,$$

which contradicts (D.3). Hence,
$$x_i > \epsilon, \quad \forall i \in [n].$$

Define three index sets
$$\mathcal{I}_1 := \{i \in [n] \mid x_i \geq 1 + \epsilon\}, \ \mathcal{I}_2 := \{i \in [n] \mid x_i \leq 1 - \epsilon\}, \ \mathcal{I}_3 := [n] \backslash (\mathcal{I}_1 \cup \mathcal{I}_2).$$

Choosing the perturbation direction $q \in \mathbb{R}^n$ to be
$$q_i = -x_i, \quad \forall i \in \mathcal{I}_1; \quad q_i = x_i, \quad \forall i \in \mathcal{I}_2; \quad q_i = 0, \quad \forall i \in \mathcal{I}_3,$$

we can calculate that

$$(\text{D.4}) \qquad \frac{1}{4} q^T g(x) = \sum_{i,j \in \mathcal{I}_1, i \neq j} -x_i x_j (x_i x_j - 1) + \sum_{i,j \in \mathcal{I}_2, i \neq j} x_i x_j (x_i x_j - 1)$$
$$+ \sum_{i \in \mathcal{I}_1, j \in \mathcal{I}_3} -x_i x_j (x_i x_j - 1) + \sum_{i \in \mathcal{I}_2, j \in \mathcal{I}_3} x_i x_j (x_i x_j - 1).$$

We then consider four sub-cases.

Case I-1. We first assume that $|\mathcal{I}_1| \geq 2$. In this case, we have

$$(\text{D.5}) \quad \sum_{i,j \in \mathcal{I}_1, i \neq j} -x_i x_j (x_i x_j - 1) \leq \sum_{i,j \in \mathcal{I}_1, i \neq j} -x_i x_j [(1 + \epsilon)^2 - 1] \leq -2\epsilon \sum_{i,j \in \mathcal{I}_1, i \neq j} x_i x_j$$
$$\leq -2\epsilon(|\mathcal{I}_1| - 1)\|x_{\mathcal{I}_1}\|_1 \leq -2\|x_{\mathcal{I}_1}\|_1 \cdot \epsilon,$$

$$\sum_{i,j \in \mathcal{I}_2, i \neq j} x_i x_j (x_i x_j - 1) = \sum_{i,j \in \mathcal{I}_2, i \neq j} -x_i \cdot [x_j - x_i x_j^2]$$
$$\leq \sum_{i,j \in \mathcal{I}_2, i \neq j} x_i \cdot [x_j - (1 - \epsilon)x_j^2]$$
$$\leq \sum_{i \in \mathcal{I}_2} x_i \max(|\mathcal{I}_2| - 1, 0) \cdot \epsilon[1 - (1 - \epsilon)\epsilon]$$
$$= -\max(|\mathcal{I}_2| - 1, 0)\|x_{\mathcal{I}_2}\|_1 \cdot \epsilon + O(n\epsilon^2),$$

$$\sum_{i \in \mathcal{I}_1, j \in \mathcal{I}_3} -x_i x_j (x_i x_j - 1) = \sum_{i \in \mathcal{I}_1, j \in \mathcal{I}_3} -\frac{1}{4}(2x_i x_j - 1)^2 + \frac{1}{4}$$
$$\leq |\mathcal{I}_1||\mathcal{I}_3| \left[-\frac{1}{4}[2(1 + \epsilon)(1 - \epsilon) - 1]^2 + \frac{1}{4} \right],$$
$$= |\mathcal{I}_1||\mathcal{I}_3| \left(\epsilon^2 - \epsilon^4 \right) = O(n^2 \epsilon^2),$$

$$\sum_{i \in \mathcal{I}_2, j \in \mathcal{I}_3} x_i x_j (x_i x_j - 1) = \sum_{i \in \mathcal{I}_2, j \in \mathcal{I}_3} \frac{1}{4}(2x_i x_j - 1)^2 - \frac{1}{4}$$
$$\leq |\mathcal{I}_2||\mathcal{I}_3| \left[\frac{1}{4}[2(1 + \epsilon)(1 - \epsilon) - 1]^2 - \frac{1}{4} \right] \leq 0.$$

Choosing ϵ to be small enough and substituting the above four estimates into (D.4), we obtain that

$$\frac{1}{4} q^T g(x) \leq -2\epsilon\|x_{\mathcal{I}_1}\|_1 - \max(|\mathcal{I}_2| - 1, 0)\epsilon\|x_{\mathcal{I}_2}\|_1 + O(n^2 \epsilon^2)$$
$$\leq -\left[\|x_{\mathcal{I}_1}\|_1 + \max(|\mathcal{I}_2| - 1, 0)\|x_{\mathcal{I}_2}\|_1 \right] \cdot \epsilon.$$

If $|\mathcal{I}_2| \geq 2$, it follows from Hölder's inequality that

$$\|g(x)\|_\infty \geq \frac{4(\|x_{\mathcal{I}_1}\|_1 + \|x_{\mathcal{I}_2}\|_1) \cdot \epsilon}{\|q\|_1} = \frac{\|x_{\mathcal{I}_1}\|_1 + \|x_{\mathcal{I}_2}\|_1}{\|x_{\mathcal{I}_1}\|_1 + \|x_{\mathcal{I}_2}\|_1} \cdot 4\epsilon = 4\epsilon.$$

which is a contradiction to (D.2). Otherwise if $|\mathcal{I}_2| \leq 1$, it also follows from Hölder's inequality that

$$\|g(x)\|_\infty \geq \frac{4\|x_{\mathcal{I}_1}\|_1 \epsilon}{\|q\|_1} = \frac{4\|x_{\mathcal{I}_1}\|_1}{\|x_{\mathcal{I}_1}\|_1 + \|x_{\mathcal{I}_2}\|_1} \cdot \epsilon \geq \frac{\|x_{\mathcal{I}_1}\|_1}{\|x_{\mathcal{I}_1}\|_1 + 1} \cdot 4\epsilon \geq 2\epsilon.$$

In summary, in this sub-case, we have

$$\|g(x)\|_\infty \geq 2\epsilon.$$

Case I-2. Now, we consider the case when $|\mathcal{I}_1| = 1$ and $|\mathcal{I}_2| \geq 2$. Assume without loss of generality that $\mathcal{I}_1 = \{1\}$. A similar calculation as (D.5) leads to

$$\frac{1}{4}q^T g(x) \leq -\max(|\mathcal{I}_2| - 1, 0)\epsilon\|x_{\mathcal{I}_2}\|_1 + O(n^2\epsilon^2) \leq -\frac{1}{2}\|x_{\mathcal{I}_2}\|_1 \cdot \epsilon.$$

If $x_1 \leq 2\epsilon^{-1}$, Hölder's inequality gives

$$\|g(x)\|_\infty \geq \frac{4\epsilon\|x_{\mathcal{I}_2}\|_1}{2\|q\|_1} = 2\epsilon \cdot \frac{\|x_{\mathcal{I}_2}\|_1}{\|x_{\mathcal{I}_1}\|_1 + \|x_{\mathcal{I}_2}\|_1} \geq 2\epsilon \cdot \frac{2\epsilon}{2\epsilon^{-1} + 2\epsilon} \geq 2\epsilon \cdot \frac{\epsilon^2}{2} = \epsilon^3.$$

Now, we assume that $x_1 > 2\epsilon^{-1}$. The first component of the gradient is

$$\frac{1}{4}g_1(x) = \sum_{j \in [n], j \neq 1} (x_j^2 x_i - x_j) \geq \sum_{j \in [n], j \neq 1} (\epsilon^2 x_i - \epsilon)$$
$$= (n-1)\epsilon^2 \cdot x_1 - (n-1)\epsilon > (n-1)\epsilon > \epsilon,$$

which contradicts (D.2). In summary, in this sub-case, we have

$$\|g(x)\|_\infty \geq \epsilon^3.$$

Case I-3. In this case, we assume $|\mathcal{I}_1| = 1$ and $|\mathcal{I}_2| \leq 1$. In addition, we assume $\mathcal{I}_1 = \{1\}$. If $x_1 \geq (1 - \epsilon)^{-1} + \epsilon$, the third estimate in (D.5) becomes

$$\sum_{j \in \mathcal{I}_3} -x_1 x_j(x_1 x_j - 1) \leq \sum_{j \in \mathcal{I}_3} -x_1(1-\epsilon)[x_1(1-\epsilon) - 1]$$
$$\leq -(1-\epsilon)^2 \epsilon x_1 \leq -\frac{1}{2}\|x_{\mathcal{I}_1}\|_1 \cdot \epsilon.$$

Then, using a similar analysis and by applying Hölder's inequality, it follows that

$$\frac{1}{4}q^T g(x) \leq -\frac{1}{2}\|x_{\mathcal{I}_1}\|_1 \epsilon \quad \text{and} \quad \|g(x)\|_\infty \geq 2\epsilon \cdot \frac{\|x_{\mathcal{I}_1}\|_1}{\|x_{\mathcal{I}_1}\|_1 + \|x_{\mathcal{I}_2}\|_1} > \epsilon.$$

Otherwise, if $x_1 < (1 - \epsilon)^{-1} + \epsilon$,

$$|x_1 - 1| < \frac{\epsilon}{1 - \epsilon} + \epsilon < 3\epsilon.$$

Hence,

$$\|x - x^0\|_1 \leq 3\epsilon + (n-1)\epsilon = (n+2)\epsilon.$$

In summary, in this sub-case, we have

$$\|g(x)\|_\infty < \epsilon/4 \quad \text{or} \quad \|x - x^0\|_1 \leq (n+2)\epsilon.$$

Case I-4. Finally, we assume $|\mathcal{I}_1| = 0$. If $|\mathcal{I}_2| \geq 2$, we can use a similar analysis as *Case I-2* to conclude that

$$\frac{1}{4}q^T g(x) \leq -\frac{1}{2}\|x_{\mathcal{I}_2}\|_1 \cdot \epsilon + O(n\epsilon^2)$$

and thus

$$\|g(x)\|_\infty \geq \epsilon.$$

Next, we consider the case when $|\mathcal{I}_2| = 1$ and we assume $\mathcal{I}_2 = \{1\}$. The fourth term in (D.5) can be estimated as

$$\sum_{i \in \mathcal{I}_2, j \in \mathcal{I}_3} x_i x_j (x_i x_j - 1) = \left(\sum_{j=2}^n x_j^2\right) x_1^2 - \left(\sum_{j=2}^n x_j\right) x_1.$$

Since $x_j \in [1 - \epsilon, 1 + \epsilon]$ for all $j \in \{2, \ldots, n\}$, it holds that

$$\frac{\sum_{j=2}^n x_j}{\sum_{j=2}^n x_j^2} \geq \frac{1}{1+\epsilon} > 1 - \epsilon.$$

Therefore,

$$\begin{aligned}
\sum_{i \in \mathcal{I}_2, j \in \mathcal{I}_3} x_i x_j (x_i x_j - 1) &= \left(\sum_{j=2}^n x_j^2\right) x_1^2 - \left(\sum_{j=2}^n x_j\right) x_1 \\
&\leq \left(\sum_{j=2}^n x_j^2\right)(1 - \epsilon)^2 - \left(\sum_{j=2}^n x_j\right)(1 - \epsilon) \\
&= \sum_{j=2}^n \left[(1 - \epsilon)^2 x_j^2 - (1 - \epsilon)x_j\right] \\
&\leq \sum_{j=2}^n \left[(1 - \epsilon)^2(1 + \epsilon)^2 - (1 - \epsilon)(1 + \epsilon)\right] \\
&\leq -(n - 1)\epsilon^2 + O(n\epsilon^3).
\end{aligned}$$

Thus, it holds that

$$\frac{1}{4}q^T g(x) \leq -(n - 1)\epsilon^2 + O(n\epsilon^3) \geq -\epsilon^2.$$

Hölder's inequality implies that

$$\|g(x)\|_\infty \geq \frac{4\epsilon^2}{\|q\|_1} = \frac{4\epsilon^2}{x_1} \geq \frac{4\epsilon^2}{1 - \epsilon} \geq 4\epsilon^2.$$

The only remaining case is when $|\mathcal{I}_2| = 0$. In this case, we have

$$x_i \in [1 - \epsilon, 1 + \epsilon], \quad \forall i \in [n].$$

Therefore, it holds that

$$\|x - x^0\|_1 \leq n\epsilon.$$

In summary, in this sub-case, we have

$$\|g(x)\|_\infty \geq 4\epsilon^2 \quad \text{or} \quad \|x - x^0\|_1 \leq n\epsilon.$$

Combining *Cases I-1* to *I-4* yields that

$$\|g(x)\|_\infty \geq \epsilon^3 \quad \text{or} \quad \|x - x^0\|_1 \leq (n + 4)\epsilon$$

in *Case I*.

Case II. For the case when $\sum_{i \in [n]} x_i < -2\epsilon$, one can obtain the same conclusions as *Case I* by the symmetry of the landscape.

Case III. We finally consider the case when $\sum_{i \in [n]} x_i \in [-2\epsilon, 2\epsilon]$. Considering the assumption (D.2), we have

$$\frac{1}{4} g_i(x) = \left(\sum_{j \in [n], j \neq i} x_j^2 + 1 \right) x_i - \sum_{j \in [n]} x_j \in [-\epsilon, \epsilon], \quad \forall i \in [n].$$

Combined with the assumption that $\sum_{i \in [n]} x_i \in [-2\epsilon, 2\epsilon]$, it follows that

$$\left(\sum_{j \in [n], j \neq i} x_j^2 + 1 \right) x_i \in [-3\epsilon, 3\epsilon].$$

Furthermore, since $\sum_{j \in [n], j \neq i} x_j^2 + 1 \geq 1$, we have

$$x_i \in [-3\epsilon, 3\epsilon], \quad \forall i \in [n].$$

We consider the descent direction $p \in \mathbb{R}^n$, where

$$p_i = 1/\sqrt{n}, \quad \forall i \in [n].$$

Then, we can calculate that

$$
\begin{aligned}
\frac{1}{4} p^T H(x) p &= \sum_{i,j \in [n] j \neq i} \left[x_j^2 p_i^2 + (2 x_i x_j - 1) p_i p_j \right] \\
&= \frac{1}{n} \sum_{i,j \in [n] j \neq i} \left[x_j^2 + (2 x_i x_j - 1) \right] \\
&= \frac{1}{n} \left[(n-1) \sum_{i \in [n]} x_i^2 + 2 \sum_{i,j \in [n], i \neq j} x_i x_j - n(n-1) \right] \\
&\leq \frac{1}{n} \left[(n-1) \cdot 9 n \epsilon^2 + 2 n(n-1) \cdot 9 \epsilon^2 - n(n-1) \right] \\
&= 27(n-1)\epsilon^2 - (n-1) \leq -n/2,
\end{aligned}
$$

where the last inequality is because ϵ is sufficiently small.

Combined *Cases I-III*, we have proved that under assumption (D.2), it holds that

$$\min\{ \|x - x^0\|_1, \|x + x^0\|_1 \} \leq (n+4)\epsilon \quad \text{or} \quad \|g(x)\|_\infty \geq \epsilon^3 \quad \text{or} \quad \lambda_{min}[H(x)] \leq -2n.$$

Letting $\epsilon := \eta/(n+4) \ll 1$, we know that the property stated in the theorem holds for problem (D.1) with

$$\beta(\eta) = \frac{\eta^3}{(n+4)^3}, \quad \gamma(\eta) = 2n.$$

In addition, we have $\eta_0 = O(1)$, $\beta(\eta) = O(n^{-3}\eta^3)$ and $\gamma(\eta) = O(n)$. Transforming back to the instance (C^0, u^0), the property stated in the theorem holds with

$$\eta_0 = O(n^{-0.5}), \quad \beta(\eta) = O(n^{-6.5}\eta^3), \quad \gamma(\eta) = O(n^{-2}).$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

D.2 Proof of Lemma 4.2

Proof Similar to Lemma 4.1, it is equivalent to prove the results for the scaled instance $\mathcal{MC}(n(n-1)\tilde{C}, n\tilde{u}^*)$. With a little abuse of notations, we use $(\tilde{C}\tilde{u}^*)$ to denote the scaled pair of parameters. Denote

$$\delta := \max \left\{ \frac{n(n-1)\epsilon}{\alpha^*}, \frac{n\epsilon}{1 - \alpha^*} \right\}.$$

Then, the condition stated in the lemma implies that

$$\tilde{C}_{ij} \in [1 - \delta, 1 + \delta], \quad \forall i, j \in [n] \quad \text{s.t.} \quad i \neq j; \quad \tilde{C}_{ii} \in [0, \delta], \quad \tilde{u}_i^* \in [1 - \delta, 1 + \delta], \quad \forall i \in [n].$$

Let $R > 0$ be a large enough constant. Suppose that $u \in \mathbb{R}^n$ is a stationary point of the instance $(\tilde{C}, \tilde{u}^*)$ such that $\|u\|_2 = R$. Denote the gradient and the Hessian matrix of the instance $\mathcal{MC}(\tilde{C}, \tilde{u}^*)$ at u as $g(u) \in \mathbb{R}^n$ and $H(u) \in \mathbb{R}^{n \times n}$, respectively. Then, it holds that

$$\text{(D.6)} \qquad \frac{1}{4} g_i(u) = \sum_{j \in [n]} \tilde{C}_{ij} u_j (u_i u_j - \tilde{u}_i^* \tilde{u}_j^*) = 0, \quad \forall i \in [n].$$

We assume without loss of generality that

$$u_1 = \max_{i \in [n]} |u_i| \geq R/\sqrt{n} > 0.$$

If $u_i = 0$ for all $i \in [n] \setminus \{1\}$, we have

$$\frac{1}{4} g_2(u) = \left(\tilde{C}_{21} u_1^2 + \sum_{j \geq 2} \tilde{C}_{2j} u_j^2 \right) u_2 - \left(\tilde{C}_{21} \tilde{u}_1^* \tilde{u}_2^* u_1 + \sum_{j \geq 2} \tilde{C}_{2j} \tilde{u}_j^* \tilde{u}_2^* u_j \right)$$
$$= -\tilde{C}_{21} \tilde{u}_1^* \tilde{u}_2^* u_1 \leq -(1 - \delta) \cdot (1 - \delta)^2 \cdot R < 0,$$

where the last inequality is in light of $\tilde{C}_{21} > 1 - \delta$ and $\tilde{u}_i^* > 1 - \delta$. This contradicts the stationarity of point x and thus

$$\sum_{j \geq 2} u_j^2 > 0.$$

Moreover, since $\tilde{C}_{1j} > 1 - \delta$ for all $j \in [n] \setminus \{1\}$, we have

$$\sum_{j \in [n]} \tilde{C}_{1j} u_j^2 \geq \sum_{j \geq 2} \tilde{C}_{1j} u_j^2 > (1 - \delta) \sum_{j \geq 2} u_j^2 > 0.$$

Similarly, for all $i \in [n] \setminus \{1\}$, it holds that

$$\sum_{j \in [n]} \tilde{C}_{ij} u_j^2 \geq \sum_{j \in [n], j \neq i} \tilde{C}_{ij} u_j^2 > (1 - \delta) \sum_{j \in [n], j \neq i} u_j^2 > (1 - \delta) u_i^2 > 0.$$

Solving (D.6) for all $i \in [n]$, we conclude that

$$\text{(D.7)} \qquad u_i = \frac{\sum_{j \in [n]} \tilde{C}_{ij} \tilde{u}_i^* \tilde{u}_j^* u_j}{\sum_{j \in [n]} \tilde{C}_{ij} u_j^2}.$$

Assuming that

$$u_1 < R - \frac{2n}{(1 - \delta) R},$$

it follows that

$$\text{(D.8)} \qquad \sum_{j \in [n]} \tilde{C}_{1j} u_j^2 > (1 - \delta) \sum_{j \geq 2} u_j^2 \geq 4n - \frac{4n^2}{(1 - \delta) R^2}.$$

In addition, we can calculate that

$$\text{(D.9)} \qquad \sum_{j \in [n]} \tilde{C}_{1j} \tilde{u}_1^* \tilde{u}_j^* u_j \leq \sum_{j \in [n]} \tilde{C}_{1j} \tilde{u}_1^* \tilde{u}_j^* |u_j|$$
$$< (1 + \delta) \cdot (1 + \delta)^2 \sum_{j \in [n]} |u_j| \leq 2 \|u\|_1 \leq 2\sqrt{n} R,$$

where the second last inequality is because δ is a sufficiently small constant. Combining inequalities (D.8)-(D.9), we have

$$u_1 = \frac{\sum_{j \in [n]} \tilde{C}_{1j} \tilde{u}_1^* \tilde{u}_j^* u_j}{\sum_{j \in [n]} \tilde{C}_{1j} u_j^2} < \frac{2\sqrt{n} R}{4n - 4n^2/[(1 - \delta)^2 R^2]}.$$

Choosing $R \geq 4n \geq 2(1 - \delta)^{-1} n$, the above inequality leads to

$$u_1 < \frac{2\sqrt{n} R}{4n - 4n^2/[(1 - \delta)^2 R^2]} < \frac{2\sqrt{n} R}{2n} = \frac{R}{\sqrt{n}},$$

which contradicts the assumption that $u_1 \geq R/\sqrt{n}$. Therefore,

$$u_1 \geq R - \frac{2n}{(1-\delta)R}.$$

Using the condition that $\|x\|_2 = R$, it holds that

$$\sum_{j \geq 2} u_j^2 \leq \frac{2n}{1-\delta} - \frac{4n^2}{(1-\delta)^2 R^2} < \frac{2n}{1-\delta}.$$

For all $i \in [n]\backslash\{1\}$, the relation (D.7) implies that

$$
\begin{aligned}
u_i &= \frac{\sum_{j \in [n]} \tilde{C}_{ij} \tilde{u}_i^* \tilde{u}_j^* u_j}{\sum_{j \in [n]} \tilde{C}_{ij} u_j^2} = \frac{\tilde{C}_{1i} \tilde{u}_i^* \tilde{u}_1^* u_1 + \sum_{j \geq 2} \tilde{C}_{ij} \tilde{u}_i^* \tilde{u}_j^* u_j}{\sum_{j \in [n]} \tilde{C}_{ij} u_j^2} \\
&\geq \frac{(1-\delta) \cdot (1-\delta)^2 (R - 2n/[(1-\delta)R]) - (1+\delta) \cdot (1+\delta)^2 \sqrt{n \cdot \sum_{j \geq 2} u_j^2}}{\sum_{j \in [n]} \tilde{C}_{ij} u_j^2} \\
&\geq \frac{(1-\delta) \cdot (1-\delta)^2 (R - 2n/[(1-\delta)R]) - (1+\delta) \cdot (1+\delta)^2 \sqrt{n \cdot 2n/(1-\delta)}}{\sum_{j \in [n]} \tilde{C}_{ij} u_j^2} \\
&\geq \frac{1/2 \cdot (R-1) - 2n\sqrt{2/(1-\delta)}}{\sum_{j \in [n]} \tilde{C}_{ij} u_j^2} \geq \frac{R/2 - 1/2 - 4n}{\sum_{j \in [n]} \tilde{C}_{ij} u_j^2} > 0,
\end{aligned}
$$

where the last inequality is due to choosing $R > 8n+1$ and the second last inequality results from the fact that δ is sufficiently small. Using the same relation, it follows that

$$
\begin{aligned}
u_i &= \frac{\sum_{j \in [n]} \tilde{C}_{ij} \tilde{u}_i^* \tilde{u}_j^* u_j}{\sum_{j \in [n]} \tilde{C}_{ij} u_j^2} \geq \frac{\tilde{C}_{1i} \tilde{u}_i^* \tilde{u}_1^* u_1}{\sum_{j \in [n]} \tilde{C}_{ij} u_j^2} \geq \frac{(1-\delta)(1-\delta)^2 u_1}{(1+\delta) \cdot R^2} \\
&\geq \frac{1}{4R^2} \cdot \left(R - \frac{2n}{(1-\delta)R} \right) \geq \frac{1}{8R},
\end{aligned}
$$

where the last inequality is due to choosing $R \geq 8n \geq 4(1-\delta)^{-1} n$. Furthermore, using the relation (D.7) with $i = 1$, we have

$$
\begin{aligned}
\sum_{j \geq 2} u_j &\geq \frac{1}{(1+\delta)(1+\delta)^2} \sum_{j \geq 2} \tilde{C}_{1j} \tilde{u}_i^* \tilde{u}_j^* u_j \\
&= \frac{1}{(1+\delta)(1+\delta)^2} \cdot u_1 \left[\sum_{j \geq 2} \tilde{C}_{1j} u_j^2 + \tilde{C}_{11}[u_1^2 - (\tilde{u}_1^*)^2] \right] \\
&\geq \frac{1}{(1+\delta)(1+\delta)^2} \cdot u_1 \Bigg[(1-\delta) \sum_{j \geq 2} u_j^2 \\
&\qquad\qquad\qquad\qquad + \tilde{C}_{11}[(R - 2n/[(1-\delta)R])^2 - (1+\delta)^2] \Bigg] \\
&\geq \frac{1-\delta}{(1+\delta)(1+\delta)^2} \cdot u_1 \left(\sum_{j \geq 2} u_j^2 \right) \\
&\geq \frac{1-\delta}{(1+\delta)(1+\delta)^2} \left(R - \frac{2n}{(1-\delta)R} \right) \cdot \sum_{j \geq 2} u_j^2 \\
&\geq \frac{1}{4} \left(R - \frac{2n}{(1-\delta)R} \right) \cdot \sum_{j \geq 2} u_j^2.
\end{aligned}
$$

Since $\sum_{j \geq 2} u_j \leq \sqrt{n(\sum_{j \geq 2} u_j^2)}$, it follows that

$$\sqrt{n \left(\sum_{j \geq 2} u_j^2 \right)} \geq \frac{1}{4} \left(R - \frac{2n}{(1-\delta)R} \right) \cdot \sum_{j \geq 2} u_j^2,$$

which further implies that

$$\sum_{j\geq 2} u_j^2 \leq \frac{16n}{(R - 2n/[(1-\delta)R])^2} \leq \frac{16n}{(R-1)^2} \leq \frac{1}{4},$$

where the last inequality is because of choosing $R \geq 1 + 8\sqrt{n}$. Now, we consider the descent direction $q \in \mathbb{R}^n$, where

$$q_1 = -u_1; \quad q_i = u_i, \quad \forall i \in [n]\backslash\{1\}.$$

Similar to the proof of Lemma 4.1, we can calculate that

$$
\begin{aligned}
\frac{1}{4}\langle g(u), q\rangle &= \sum_{i,j\geq 2, i\neq j} \tilde{C}_{ij} u_i u_j \left(u_i u_j - \tilde{u}_i^* \tilde{u}_j^*\right) - \tilde{C}_{11} u_1^2 [u_1^2 - (\tilde{u}_1^*)^2] \\
&\qquad\qquad\qquad + \sum_{i\geq 2} \tilde{C}_{ii} u_i^2 [u_i^2 - (\tilde{u}_i^*)^2] \\
&\leq \sum_{i,j\geq 2, i\neq j} \tilde{C}_{ij} u_i u_j \left(u_i u_j - \tilde{u}_i^* \tilde{u}_j^*\right) + \sum_{i\geq 2} \tilde{C}_{ii} u_i^2 [u_i^2 - (\tilde{u}_i^*)^2] \\
&\leq \sum_{i,j\geq 2, i\neq j} \tilde{C}_{ij} u_i u_j \left[1/4 - (1-\delta)^2\right] \\
&\qquad\qquad\qquad + \sum_{i\geq 2} \tilde{C}_{ii} u_i^2 [1/4 - (1-\delta)^2] \\
&\leq \sum_{i,j\geq 2, i\neq j} (1-\delta) \cdot (8R)^{-2} \cdot (1/4 - 1/2) \\
&\qquad\qquad\qquad + \sum_{i\geq 2} \delta \cdot (8R)^{-2} \cdot (1/4 - 1/2) < 0,
\end{aligned}
$$

which contradicts the assumption that x is a stationary point. Therefore, the above analysis implies that the instance $(\tilde{C}, \tilde{u}^*)$ has no stationary point in the region $\{u \in \mathbb{R}^n \mid \|u\|_2 > 8n + 1\}$.

Now, We focus on the compact region $\{u \in \mathbb{R}^n \mid \|u\|_2 \leq 8n + 1\}$. Since the gradient and the Hessian matrix are continuous functions of (C, u^*), the ℓ_∞-norm of the gradient and the eigenvalues of the Hessian matrix are also continuous functions of (C, u^*). Intuitively, a small perturbation to (C, u^*) would not significantly change the norms of the gradient and the Hessian matrix. Thus, the strict-saddle property still holds after a small perturbation. More rigorously, let $(C^0, u^0) \in \mathcal{M}$ and $\eta \in (0, \eta_0]$. In the region

$$\mathcal{R}_\eta := \{u \in \mathbb{R}^n \mid \|u\|_2 \leq 8n + 1, \|u - u^0\|_1 \geq \eta, \|u + u^0\|_1 \geq \eta\},$$

at least one of the following properties holds:

$$\|\nabla g(u; C^0, u^0)\|_\infty \geq \beta(\eta), \quad \lambda_{min}[\nabla^2 g(u; C^0, u^0)] \leq -\gamma(\eta).$$

Since $\mathcal{R}_\eta$ is a compact set and we constrain (C, u^*) by $\|C\|_1 = 1$ and $\|u^*\|_1 = 1$, the functions

$$\|\nabla g(u; C, u^*)\|_\infty \quad \text{and} \quad \lambda_{min}[\nabla^2 g(x; C, u^*)]$$

are Lipschitz continuous in (C, u^*). Suppose that the Lipschitz constants are L_g and L_H under the weighted ℓ_1-norm, namely

$$
\begin{aligned}
\Big| \|\nabla g(u; C, u^*)\|_\infty &- \|\nabla g(u; \tilde{C}, \tilde{u}^*)\|_\infty \Big| \\
&\leq L_g \left[\alpha^* \|\tilde{C} - C\|_1 + (1-\alpha^*) \|\tilde{u}^* - u^*\|_1\right], \\
\Big| \lambda_{min}[\nabla^2 g(u; C, u^*)] &- \lambda_{min}[\nabla^2 g(u; \tilde{C}, \tilde{u}^*)] \Big| \\
&\leq L_H \left[\alpha^* \|\tilde{C} - C\|_1 + (1-\alpha^*) \|\tilde{u}^* - u^*\|_1\right], \\
&\forall x \in \mathcal{R}_\eta, \ (C, u^*) \quad \text{s.t.} \quad \|C\|_1 = \|u^*\|_1 = 1.
\end{aligned}
$$

Let

$$\epsilon := \min\left\{\frac{\beta(\eta)}{2L_g}, \frac{\gamma(\eta)}{2L_H}\right\}.$$

Then, for every pair $(\tilde{C}, \tilde{u}^*)$ satisfying

$$\alpha^* \|\tilde{C} - C^0\|_1 + (1 - \alpha^*)\|\tilde{u}^* - u^0\|_1 < \epsilon,$$

at least one of the following properties holds for all $x \in \mathcal{R}_\eta$:

$$\|\nabla g(u; \tilde{C}, \tilde{u}^*)\|_\infty \geq \beta(\eta)/2, \quad \lambda_{min}[\nabla^2(u; \tilde{C}, \tilde{u}^*)] \leq -\gamma(\eta)/2.$$

This implies that the strict-saddle property holds for the the perturbed instance $\mathcal{MC}(\tilde{C}, \tilde{u}^*)$. Letting $\eta \to 0$, it follows that $\pm\tilde{u}^*$ are the only points satisfying the second-order necessary optimality conditions, and thus $\mathcal{MC}(\tilde{C}, \tilde{u}^*)$ does not have SSCPs. $\qquad\square$