Journal of Computational Mathematics
Vol.38, No.4, 2020, 638–660.
http://www.global-sci.org/jcm
doi:10.4208/jcm.1902-m2018-0109

SOLVING SYSTEMS OF QUADRATIC EQUATIONS VIA EXPONENTIAL-TYPE GRADIENT DESCENT ALGORITHM*

Meng Huang and Zhiqiang Xu
LSEC, Inst. Comp. Math., Academy of Mathematics and System Science,
Chinese Academy of Sciences, Beijing 100190, China;
School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
Email: hm@lsec.cc.ac.cn, xuzq@lsec.cc.ac.cn

Abstract

We consider the rank minimization problem from quadratic measurements, i.e., recovering a rank r matrix $X \in \mathbb{R}^{n\times r}$ from m scalar measurements $y_i = a_i^\top XX^\top a_i,\ a_i \in \mathbb{R}^n,\ i = 1,\ldots,m$. Such problem arises in a variety of applications such as quadratic regression and quantum state tomography. We present a novel algorithm, which is termed *exponential-type gradient descent algorithm*, to minimize a non-convex objective function $f(U) = \frac{1}{4m}\sum_{i=1}^m (y_i - a_i^\top UU^\top a_i)^2$. This algorithm starts with a careful initialization, and then refines this initial guess by iteratively applying exponential-type gradient descent. Particularly, we can obtain a good initial guess of X as long as the number of Gaussian random measurements is $O(nr)$, and our iteration algorithm can converge linearly to the true X (up to an orthogonal matrix) with $m = O\left(nr\log(cr)\right)$ Gaussian random measurements.

Mathematics subject classification: 90C26, 94A15.
Key words: Low-rank matrix recovery, Non-convex optimization, Phase retrieval.

1. Introduction

1.1. Problem setup.

Let $X \in \mathbb{R}^{n\times r}$ be a fixed and unknown matrix with $\mathrm{rank}(X) = r$, and our aim is to recover X from given quadratic measurements, i.e.,

$$\text{find} \quad X \in \mathbb{R}^{n\times r}, \quad \text{s.t.} \quad y_i = a_i^\top XX^\top a_i = \|a_i^\top X\|_2^2, \qquad i = 1,\ldots,m, \tag{1.1}$$

where $a_i = (a_{i,1},\ldots,a_{i,n}) \in \mathbb{R}^n$. This problem is raised in many emerging applications of science and engineering, such as covariance sketching, quantum state tomography and high dimensional data streams [7,16,17]. A simple observation is that $a_i^\top XX^\top a_i = a_i^\top XOO^\top X^\top a_i$ where $O \in \mathbb{R}^{r\times r}$ is an orthogonal matrix. We can only hope to recover X up to a right orthogonal matrix. There exists an orthogonal matrix $O^* \in \mathbb{R}^{r\times r}$ such that XO^* has orthogonal column vectors. Hence, throughout the paper we can assume that X has orthogonal column vectors.

To recover X from given measurements (1.1), we consider the following optimization problem:

$$\min_{U\in\mathbb{R}^{n\times r}} f(U) = \frac{1}{4m}\sum_{i=1}^m (y_i - \|a_i^\top U\|_2^2)^2. \tag{1.2}$$

The aim of this paper is to develop algorithms to solve (1.2).

* Received June 1, 2018 / Revised version received January 21, 2019 / Accepted February 19, 2019 / Published online July 25, 2019 /

1.2. Related work

1.2.1. Low rank matrix recovery

Rank minimization problem is a direct generalization of compressed sensing [15, 22]. For the general rank minimization problem, it aims to reconstruct a low rank matrix $Q \in \mathbb{R}^{n\times n}$ from incomplete measurements, which can be formulated as the following programming

$$\begin{aligned} \min_{Z\in\mathbb{R}^{n\times n}} \quad & \operatorname{rank}(Z) \\ \text{subject to} \quad & \operatorname{tr}(A_i Z) = y_i, \quad i = 1,\ldots,m, \end{aligned} \tag{1.3}$$

where $y_i = \operatorname{tr}(A_i Q), A_i \in \mathbb{R}^{n\times n}, i = 1,\ldots,m$. In [26], Xu has proved that in order to guarantee the solution of (1.3) is Q where $Q \in \mathbb{C}^{n\times n}$ and $\operatorname{rank}(Q) \le r$, the minimal measurement number m is $4nr - 4r^2$. Since (1.3) is non-convex, it is challenging to solve it [18]. However, under a certain restricted isometry property (RIP), this problem can be relaxed to a nuclear norm minimization problem which is a convex programming and can be solved efficiently [4, 22].

Noting that $M := XX^\top$ is a low rank matrix, we can recast (1.1) as a rank minimization problem. This means that we can use the nuclear norm minimization to recover the matrix M and hence X:

$$\begin{aligned} \min_{Z\in\mathcal{H}_n} \quad & \|Z\|_* \\ \text{subject to} \quad & \operatorname{tr}(A_i Z) = y_i, \quad i = 1,\ldots,m, \end{aligned} \tag{1.4}$$

where $\mathcal{H}_n := \{Q \in \mathbb{R}^{n\times n} : Q = Q^\top\}$ and $A_i = a_i a_i^*$. Problem (1.4) was studied in [7, 16] with proving that $m \ge Cnr$ Gaussian measurements are sufficient to recover the unknown matrix $M = XX^\top$ exactly. In [21], Rauhut and Terstiege also consider the case where the measurement vectors $a_i, i = 1,\ldots,m$ are from a tight frame.

1.2.2. Phase retrieval

Under the setting of $r = 1$, the (1.1) is reduced to phase retrieval problem. Phase retrieval is to recover an unknown vector from the magnitude of measurements, which means to recover a signal $x \in \mathbb{H}^n$ from measurements

$$y_i = |\langle a_i, x\rangle|^2, \quad i = 1,\ldots,m, \tag{1.5}$$

where $a_i \in \mathbb{H}^n$ ($\mathbb{H} = \mathbb{C}$ or $\mathbb{R}$) are sampling vectors. This problem is raised in many imaging applications due to the limitations of optical sensors which can only record intensity information, such as X-ray crystallography [14, 19], astronomy [11], diffraction imaging [13, 24]. It has been proved that $m \ge 4n - 4$ Gaussian measurements are sufficient to recover the unknown vector up to a global phase [8]. In recent years, several different algorithms have been proposed to solve it [1, 2, 9, 10, 20]. In [3], Candès et al. design Wirtinger flow algorithm for phase retrieval which solves the following non-convex optimization problem

$$\min_{u\in\mathbb{C}^n} \frac{1}{4m}\sum_{i=1}^{m}(y_i - |a_i^* u|^2)^2 \tag{1.6}$$

and prove that the algorithm converges to the true signal up to a global phase with high probability provided the measurement vectors are $m = O(n\log n)$ Gaussian measurements.

Following the work of [3], Chen and Candès [6] propose a modified gradient method which is called *Truncated Wirtinger Flow*, and it removes the additional logarithmic factor in the number of measurements m. In [12], Gao and Xu propose a Gauss-Newton algorithm to solve (1.6) and they prove that, for the real signal, the algorithm can converge to the global optimal solution quadratically with $O(n\log n)$ measurements.

1.3. Our contribution

In [23, 27], one designed algorithms for solving (1.2). In order to guarantee convergence to the global optimal solution, the algorithm in [23] requires that $m \geq C\|X\|_F^8\lambda_r^{-4}nr^2\log^2 n$, while the algorithm in [27] needs $m = O(r^3\kappa^2 n\log n)$, where κ denotes the condition number of $XX^\top$. In contrast to those algorithms, we aim to reduce the sampling complexity with removing the additional logarithmic factor on n. In this paper, we propose a novel algorithm and call it *exponential-type gradient descent algorithm*. For the initialization, we give a tighter initial guess through a careful truncated skill; and for iteration update step, we add a moderate bounded exponential-type function to the classical gradient. Particularly, we show the followings all hold with high probability:

- We present a spectral initialization method which obtains a good initial guess provided $m \geq C\sigma_r^{-2}\|X\|_F^4 nr$ and $a_i, i = 1, \dots, m$ are Gaussian random vectors, where σ_r, σ_1 are the smallest and the largest nonzero eigenvalues of the positive semidefinite matrix $XX^\top$.
- Starting from our initial guess, we refine the initial estimation by iteratively applying a novel gradient update rule. If $m \geq C\sigma_r^{-2}\|X\|_F^4 nr\log(cr\|X\|_F^2/\sigma_r)$, then our algorithm linearly converges to a global minimizer X, up to a right orthogonal matrix. More importantly, the step size in our algorithm is independent with the dimension n.

1.4. Organization

The paper is organized as follows. First, we introduce some notations and lemmas in Section 2. In Section 3, we introduce the exponential-type gradient descent algorithm for solving (1.2). We study the convergence property of the new algorithm in Section 4. In Section 5, we introduce the main idea for proving the results which are given in Section 4. Numerical experiments are made in Section 6. At last, most of the detailed proofs are given in the Appendix.

2. Preliminaries

2.1. Notations

Throughout the paper, we assume that $X = (x_1, \dots, x_r) \in \mathbb{R}^{n\times r}$ has orthogonal columns. Without loss of generality, we assume that $\|x_1\|_2 \geq \|x_2\|_2 \geq \cdots \geq \|x_r\|_2$. We use the Gaussian random vectors $a_i \in \mathbb{R}^n$, $i = 1, \dots, m$ as the measurement vectors and obtain $y_i = a_i^\top XX^\top a_i$, $i = 1, \dots, m$. Here we say the sampling vectors are the Gaussian random measurements if $a_i \in \mathbb{R}^n$ are i.i.d. $\mathcal{N}(0, I)$ random variables. As we have the entire manifold solutions given by $\mathcal{X} := \{XO : O \in \mathcal{O}(r)\}$, where $\mathcal{O}(r)$ is the set of $r\times r$ orthogonal matrices, we define the distance between a matrix $U \in \mathbb{R}^{n\times r}$ and X as

$$d(U) \;:=\; \min_{O\in\mathcal{O}(r)} \|XO - U\|_F. \tag{2.1}$$

To state conveniently, we assume that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 \tag{2.2}$$

are the nonzero eigenvalues of the matrix $XX^\top$.

2.2. Lemmas

We now introduce some lemmas which will be used in our paper. First, we recall a result about random matrix with non-isotropic sub-gaussian rows [25, Eq. (4.22)].

Lemma 2.1 ([25, Eq. (4.22)]). *Let A be an $N \times n$ matrix whose rows are A_i, and assume that $\Sigma^{-1/2}A_i$ are isotropic sub-gaussian random vectors, and let K be the maximum of their sub-gaussian norms. Then for every $t \geq 0$, the following inequality holds with probability at least $1 - 2\exp(-ct^2)$:*

$$\|\frac{1}{N}A^*A - \Sigma\|_2 \leq \max(\delta, \delta^2)\|\Sigma\|_2 \qquad \textit{where} \quad \delta = C\sqrt{\frac{n}{N}} + \frac{t}{\sqrt{N}}.$$

Here C, c are constants.

The next result is Bernstein-type inequality about sub-exponential random variables [25, Theorem 2.8.2].

Lemma 2.2 ([25, Theorem 2.8.2]). *Let $X_1, \ldots, X_N$ be independent centered sub-exponential random variables and $K = \max_i \|X_i\|_{\psi_1}$. Then for every $a = (a_1, \ldots, a_N) \in \mathbb{R}^N$ and every $t \geq 0$, we have*

$$\mathbb{P}\Big\{|\sum_{i=1}^N a_i X_i| \geq t\Big\} \leq 2\exp\Big[-c\min\Big(\frac{t^2}{K^2\|a\|_2^2}, \frac{t}{K\|a\|_\infty}\Big)\Big],$$

where $c > 0$ is an absolute constant.

Lemma 2.3. *For any $\delta > 0$, assume that $m \geq 16\delta^{-2}n$ and $a_i, i = 1, \ldots, m$ are Gaussian random vectors. Then for any positive semidefinite matrices $M \in \mathbb{R}^{n\times n}$,*

$$(1-\delta)\|M\|_* \leq \frac{1}{m}\sum_{i=1}^m a_i^\top M a_i \leq (1+\delta)\|M\|_*$$

holds on an event E_δ of probability at least $1 - 2\exp(-m\epsilon^2/2)$, where $\delta/4 = \epsilon^2 + \epsilon$ and the norm $\|\cdot\|_*$ denotes the nuclear norm of a matrix. In particular, the right inequality holds for all matrices.

Proof. The first part of this lemma is a direct consequence of Lemma 3.1 in [5]. Hence, we only need to prove that the right inequality holds for all matrices. We assume the rank of matrix M is r. Then by the singular-value decomposition, we can write $M = \sum_{j=1}^r \sigma_j u_j v_j^\top$, where u_j, v_j are unit vectors. It implies that we just need to show

$$\frac{1}{m}\sum_{i=1}^m (a_i^\top u)(a_i^\top v) \leq 1 + \delta$$

holds for any fixed unit vectors u, v. Indeed, if we denote $A := (a_1, \ldots, a_m)^\top$, then

$$\begin{aligned}&\sum_{i=1}^m (a_i^\top u)(a_i^\top v)\\ \le\ & \frac{1}{2}\sum_{i=1}^m (a_i^\top u)^2 + \frac{1}{2}\sum_{i=1}^m (a_i^\top v)^2\\ =\ & \frac{1}{2}(\|Au\|_2^2 + \|Av\|_2^2) \le \sigma_{\max}^2(A),\end{aligned}$$

where $\sigma_{\max}^2(A)$ is the maximum singular value of A. From the well known deviations bounds concerning the singular values of Gaussian random matrices, i.e.,

$$\mathbb{P}(\sigma_{\max}(A) \ge \sqrt{m} + \sqrt{n} + t) \le \exp(-t^2/2),$$

we arrive the conclusion if we take $m \ge \epsilon^{-2} n$ and $t = \sqrt{m}\epsilon$.

3. Exponential-type Gradient Descent Algorithm

Our aim is to recover a matrix $X \in \mathbb{R}^{n\times r}$ (up to right multiplication by an orthogonal matrix) from quadratic measurements

$$y_i = \|a_i^\top X\|_2^2, \qquad i = 1, \ldots, m$$

by solving the non-convex optimization problem

$$\min_{U\in\mathbb{R}^{n\times r}} f(U) = \frac{1}{4m}\sum_{i=1}^m (y_i - \|a_i^\top U\|_2^2)^2. \tag{3.1}$$

In this section, we will introduce an exponential-type gradient descent algorithm for solving (3.1).

3.1. Spectral Initialization

The first step of our algorithm is to choose a good initial guess. In [23], Sanghavi, Ward and White choose $U_0 = Z\Lambda^{1/2}$ as the initial guess, where the columns of $Z \in \mathbb{R}^{n\times r}$ are the normalized eigenvectors corresponding to the r largest eigenvalues $\lambda_1 \ge \cdots \ge \lambda_r$ of the matrix $Y = \frac{1}{2m}\sum_{i=1}^m y_i a_i a_i^\top$ and the diagonal matrix $\Lambda = \operatorname{diag}(\Lambda_1, \ldots, \Lambda_r)$ is given by $\Lambda_i = \lambda_i - \lambda_{r+1}$. To guarantee the convergence of the iterative method, the initialization method introduced in [23] requires $O(nr^2 \log^2 n)$ measurements [23]. Motivated by the methods for choosing the initial guess in [6] and [23], we introduce a novel initialization method which is stated in Algorithm 3.1. We prove that the new method just need $O(nr)$ measurements to obtain the same accuracy as the method suggested in [23].

Algorithm 3.1. Initialization

Input: Measurements $y_i = \|a_i^\top X\|^2, i = 1, \dots, m$, where a_i are Gaussian random vectors; parameter $\alpha_y > 0$.

Define $U_0 = U\Sigma^{1/2}$, where the columns of U are the normalized eigenvectors corresponding to the r largest eigenvalues $\lambda_1 \geq \cdots \geq \lambda_r$ of the matrix

$$Y = \frac{1}{m}\sum_{i=1}^{m} y_i a_i a_i^\top \mathbb{1}_{\{y_i \leq \frac{\alpha_y}{m}\sum_{k=1}^{m} y_k\}}$$

and the diagonal matrix Σ is given by

$$\Sigma_{i,i} = \frac{1}{2}(\lambda_i - \lambda_{r+1}).$$

Output: Initial guess U_0.

In our analysis, we require that the parameter α_y in Algorithm 1 satisfies $\alpha_y \geq C\sqrt{\log(c\kappa r)}$, where κ is the ratio of the largest to the smallest nonzero eigenvalues of matrix $XX^\top$ and C, c are universal constants. It means that the choice of α_y only depends on the condition number κ and the rank r of X.

3.2. Exponential-type Gradient Descent

The next step of our algorithm is to refine the initial guess by an update rule to search the global optimal solution. In [23], Sanghavi, Ward and White iteratively update U via gradient descent and they also prove the gradient descent method converges to the global optimal solution provided $m \geq Cnr\log^2 n$. We next introduce an exponential-type gradient descent update rule.

For $k = 0, 1, \dots$, we take the iteration step as

$$U_{k+1} = U_k - \mu\nabla f_{\text{ex}}(U_k), \tag{3.2}$$

where $\nabla f_{\text{ex}}(\cdot)$ denotes the exponential-type gradient given by

$$\nabla f_{\text{ex}}(U) = \frac{1}{m}\sum_{i=1}^{m}(a_i^\top UU^\top a_i - a_i^\top XX^\top a_i)a_i a_i^\top U \cdot \exp\Big(-\frac{my_i}{\alpha\sum_{k=1}^{m} y_k}\Big), \tag{3.3}$$

where $\alpha > 0$. We state our algorithm as follows:

Algorithm 3.2. Exponential-type Gradient Descent Algorithm

Input: Measurement vectors: $a_i \in \mathbb{R}^n, i = 1, \dots, m$; Observations: $y \in \mathbb{R}^m$; Parameter α; Step size μ; $\epsilon > 0$

1: Set $T := c\log\frac{1}{\epsilon}$, where c is a sufficient large constant.

2: Use Algorithm 1 to compute an initial guess U_0 .

3: or $k = 0, 1, 2, \ldots, T-1$ do

$$\begin{aligned} U_{k+1} &= U_k - \mu \nabla f_{\text{ex}}(U_k) \\ &= U_k - \frac{\mu}{m} \sum_{i=1}^m (a_i^\top U U^\top a_i - y_i) a_i a_i^\top U \cdot \exp\Big(- \frac{m y_i}{\alpha \sum_{k=1}^m y_k} \Big) \end{aligned}$$

4: End for

Output: The matrix U_T.

Remark 3.1. There is a parameter α in Algorithm 2. Throughout this paper, we select the parameter $\alpha \geq 20$. Numerical experiments in Section 6 show that the algorithm's performance is not sensitive to the selection of α.

4. Main results

In this section we present our main results which give the theoretical guarantee of Algorithm 2. We first study Algorithm 1 with showing that our initial guess U_0 is not far from $\{XO : O \in \mathcal{O}(r)\}$.

Theorem 4.1. *Suppose that* $m \geq c_0 \sigma_r^{-2} \|X\|_F^4 nr$ *and*

$$y_i = a_i^\top X X^\top a_i = \|a_i^\top X\|_2^2, \quad i = 1, \ldots, m$$

where $a_i \in \mathbb{R}^n$ *is the Gaussian random vector. Let* U_0 *be the output of Algorithm* 3.1 *with* $\alpha_y \geq C\sqrt{\log(c\kappa r)}$, *where* $\kappa = \sigma_1/\sigma_r$ *denotes the ratio of the largest to the smallest nonzero eigenvalues of the matrix* $XX^\top$. *Then with probability at least* $1 - 6\exp(-\Omega(n))$ *we have*

$$d(U_0) \leq \sqrt{\frac{\sigma_r}{8}},$$

where c, c_0 *and* C *are absolute constants, and* $d(U_0)$ *is defined as*

$$d(U_0) := \min_{O \in \mathcal{O}(r)} \|XO - U_0\|_F.$$

We next consider the convergence property of Algorithm 2.

Theorem 4.2. *Suppose that* $m \geq c_0 \sigma_r^{-2} \|X\|_F^4 nr \log(c_1 r \|X\|_F^2/\sigma_r)$ *and*

$$y_i = a_i^\top X X^\top a_i = \|a_i^\top X\|_2^2, \quad i = 1, \ldots, m$$

where $a_i \in \mathbb{R}^n$ *is the Gaussian random vector. Then the following holds with probability at least* $1 - C\exp(-\Omega(n))$. *For all* $U_k \in \mathbb{R}^{n \times r}$ *satisfies* $d(U_k) \leq \sqrt{\sigma_r/8}$, *the* U_{k+1} *defined by the update rule* (3.2) *with the step size* $\mu \leq \frac{\sigma_r^3}{c_2 \sigma_1 \|X\|_F^6}$ *satisfies*

$$d(U_{k+1}) \leq \Big(1 - \rho_0\Big)^{1/2} d(U_k), \tag{4.1}$$

where $\rho_0 = 2\mu\sigma_r/7$.

Combining Theorems 4.1 and 4.2, we can obtain the following corollary which shows that Algorithm 2 is convergent with high probability provided $m \geq Cnr\log(cr)$.

Corollary 4.1. *Suppose that* $m \geq c_0\sigma_r^{-2}\|X\|_F^4 nr\log(c_1 r\|X\|_F^2/\sigma_r)$ *and* $y_i = a_i^\top XX^\top a_i = \|a_i^\top X\|_2^2$, $i = 1,\dots,m$ *where* $a_i \in \mathbb{R}^n$ *is the Gaussian random vector. Suppose that* ϵ *is an arbitrary constant within range* $(0, \sqrt{\sigma_r/8})$. *Then with probability at least* $1 - C\exp(-\Omega(n))$, *Algorithm* 2 *outputs* U_T *satisfying*

$$d(U_T) \ \leq \ \epsilon$$

provided the step size $\mu \leq \frac{\sigma_r^3}{c_2\sigma_1\|X\|_F^6}$ *where* $T \geq \log\frac{\sigma_r}{8\epsilon^2}\log\frac{1}{1-\rho_0}$ *and* $\rho_0 = \frac{2\mu\sigma_r}{7}$.

Proof. According to Theorem 4.1, with probability at least $1 - 6\exp(-\Omega(n))$ we have

$$d(U_0) \leq \sqrt{\frac{\sigma_r}{8}}.$$

From the iterative inequality (4.1) in Theorem 4.2, we obtain that

$$\begin{aligned} d(U_T) &\leq \Big(1-\rho_0\Big)^{1/2} d(U_{T-1}) \leq \Big(1-\rho_0\Big)^{T/2} d(U_0) \\ &\leq \sqrt{\frac{\sigma_r}{8}}\Big(1-\rho_0\Big)^{T/2} \leq \epsilon, \end{aligned}$$

which holds with probability at least $1 - C\exp(-\Omega(n))$. □

Remark 4.1. According to Theorem 4.2, to guarantee Algorithm 2 converges to the true matrix, we require that the step size

$$\mu \ \leq \ \sigma_r^3/(C\sigma_1\|X\|_F^6). \tag{4.2}$$

Noting that $\|X\|_F^4 = (\sigma_1 + \cdots + \sigma_r)^2 \leq r^2\sigma_1^2$, we have $\sigma_r^3/(C\sigma_1\|X\|_F^6) \geq 1/(C\kappa^3 r^2\|X\|_F^2)$ which implies that

$$\mu \ \leq \ 1/(C\kappa^3 r^2\|X\|_F^2) \tag{4.3}$$

is enough to guarantee (4.2) holds. Recall that the algorithms in [23] and [27] require that $\mu \leq (1/Cn^4\log^4(nr)\|X\|_F^2)$ and $\mu \leq C/(\kappa n\|X\|_F^2)$, respectively. Comparing with the step size in [23] and [27], our step size is independent with the matrix dimension n.

5. The Proof of the Main Results

In this section we give the proof of the main results. To state conveniently, for $U \in \mathbb{R}^{n\times r}$, we set

$$\bar{X} := \bar{X}_U := \operatorname*{argmin}_{Z\in\mathcal{X}} \|U - Z\|_F, \tag{5.1}$$

where $\mathcal{X} := \{XO : O \in \mathcal{O}(r)\}$, and $\mathcal{O}(r)$ is the set of $r\times r$ orthogonal matrices.

Motivated by the results in [3], we next give the definition of the regularity condition. Under this condition, we shall prove that our algorithm converges linearly to the true matrix X if the initial guess is not far from it.

Definition 5.1 (Regularity Condition) *We say that the function* f *satisfies the regularity condition* $RC(\nu,\lambda,\varepsilon)$ *if there exist constants* ν,λ *such that for all matrices* $U \in \mathbb{R}^{n\times r}$ *satisfying* $d(U) \leq \varepsilon$ *we have*

$$\langle \nabla f_{\rm ex}(U), U - \bar{X}\rangle \geq \frac{1}{\nu}\sigma_r\|U - \bar{X}\|_F^2 + \frac{1}{\lambda\|X\|_F^2}\|\nabla f_{\rm ex}(U)\|_F^2,$$

where $\nabla f_{\rm ex}(\cdot)$ *is defined in* (3.3) *and* $\bar{X}$ *is defined in* (5.1).

Under the assumption of f satisfying the regularity condition, the next lemma shows the performance of the update rule.

Lemma 5.1. *Assume that the function f satisfies the regularity condition $RC(\nu,\lambda,\varepsilon)$ and $d(U_k)\le\varepsilon$. If we take the step size $\mu\le\min\left(\frac{\nu}{2\sigma_r},\frac{2}{\lambda\|X\|_F^2}\right)$, then $U_{k+1}=U_k-\mu\nabla f_{\rm ex}(U_k)$ satisfies*

$$d(U_{k+1}) \;\le\; \sqrt{1-\frac{2\mu\sigma_r}{\nu}}d(U_k).$$

Proof. To state conveniently, we set

$$\bar{X}_k := \operatorname*{argmin}_{Z\in\mathcal{X}} \|U_k-Z\|_F. \tag{5.2}$$

Under the regularity condition $RC(\nu,\lambda,\varepsilon)$, we have

$$\begin{aligned}
&d(U_{k+1})^2 \le \|U_k-\bar{X}_k-\mu\nabla f_{\rm ex}(U_k)\|_F^2\\
&=\|U_k-\bar{X}_k\|_F^2-2\mu\langle\nabla f_{\rm ex}(U_k),U-\bar{X}_k\rangle+\mu^2\|\nabla f_{\rm ex}(U_k)\|_F^2\\
&\le\|U_k-\bar{X}_k\|_F^2-2\mu\Big(\frac{1}{\nu}\sigma_r\|U_k-\bar{X}\|_F^2+\frac{1}{\lambda\|X\|_F^2}\|\nabla f_{\rm ex}(U_k)\|_F^2\Big)+\mu^2\|\nabla f_{\rm ex}(U_k)\|_F^2\\
&=\left(1-\frac{2\mu\sigma_r}{\nu}\right)\|U_k-\bar{X}_k\|_F^2+\mu(\mu-\frac{2}{\lambda\|X\|_F^2})\|\nabla f_{\rm ex}(U_k)\|_F^2\\
&\le\left(1-\frac{2\mu\sigma_r}{\nu}\right)d(U_k)^2,
\end{aligned}\tag{5.3}$$

where the last inequality follows from $\mu\le\frac{2}{\lambda\|X\|_F^2}$. □

Based on Lemma 5.1, the key point to prove Theorem 4.2 is to show that the function f satisfies the regularity condition with high probability. The next lemma shows that f satisfies the regularity condition provided $m\ge c_0\sigma_r^{-2}\|X\|_F^4nr\log(c_1r\|X\|_F^2/\sigma_r)$.

Lemma 5.2. *Suppose $m\ge c_0\sigma_r^{-2}\|X\|_F^4nr\log(c_1r\|X\|_F^2/\sigma_r)$ and f is defined as (1.2). Then f satisfies the regularity condition $RC\Big(7,\frac{250\alpha^2\sigma_1\|X\|_F^4}{\sigma_r^3},\sqrt{\frac{1}{8}\sigma_r}\Big)$ with probability at least $1-C\exp(-\Omega(n))$, where α is the constant in $\nabla f_{\rm ex}$ and C,c_0,c_1 are universal constants.*

We next state the proof of Theorem 4.2.

Proof of Theorem 4.2. According to Lemma 5.2, if $m\ge c_0\sigma_r^{-2}\|X\|_F^4nr\log(c_1r\|X\|_F^2/\sigma_r)$, then f satisfies the regularity condition with $\nu=7$, $\lambda=250\alpha^2\sigma_1\|X\|_F^4/\sigma_r^3$ and $\varepsilon=\sqrt{\sigma_r/8}$ with probability at least $1-C\exp(-\Omega(n))$. Noting that $d(U_k)\le\sqrt{\frac{1}{8}\sigma_r}$, Lemma 5.1 implies that

$$d(U_{k+1}) \;\le\; \sqrt{1-\frac{2\mu\sigma_r}{\nu}}d(U_k)=\Big(1-\frac{2\mu\sigma_r}{7}\Big)^{1/2}d(U_k)$$

provided that the step size satisfies

$$\mu\le\min\left(\frac{\nu}{2\sigma_r},\frac{2}{\lambda\|X\|_F^2}\right)=\frac{\sigma_r^3}{125\alpha^2\sigma_1\|X\|_F^6}=\frac{\sigma_r^3}{c_2\sigma_1\|X\|_F^6}.$$

□

We remain to prove Lemma 5.2. To this end, we introduce one proposition and the full details can be found in the appendix.

Proposition 5.1. *Assume that* $\|X\|_F = 1$ *and that* $m \ge c_0\sigma_r^{-2}nr\log(c_1r/\sigma_r)$. *Then with probability at least* $1 - C\exp(-\Omega(n))$, *the followings hold for all matrices* $U \in \mathbb{R}^{n\times r}$ *satisfying* $d(U) \le \sqrt{\frac{\sigma_r}{8}}$*:*

$$(a) \qquad \langle \nabla f_{\rm ex}(U), H\rangle \ge 0.166\sigma_r\|H\|_F^2 + 0.78\left(\mathrm{tr}^2(H^\top\bar{X}) + \|H^\top\bar{X}\|_F^2\right), \tag{5.4}$$

$$(b) \qquad \frac{\sigma_r^2\|\nabla f_{\rm ex}(U)\|_F^2}{3\alpha^2\left(\|H\|_F^2+\|X\|_F^2\right)} \le 1.223\sigma_1\|H\|_F^2 + \mathrm{tr}^2(H^\top\bar{X}) + \|H^\top\bar{X}\|_F^2,, \tag{5.5}$$

where $H = U - \bar{X}$ *and* $\bar{X}$ *is defined in* (5.1).

Now, we can give the proof of Lemma 5.2.

Proof of Lemma 5.2. In order to prove Lemma 5.2, we only need to consider the case where $\|X\|_F = 1$. For any $0 < \gamma < 1$, multiplying $\gamma\sigma_r/\sigma_1$ on both sides of (5.5) we have

$$\begin{aligned}&\frac{\gamma\sigma_r^3\|\nabla f_{\rm ex}(U)\|_F^2}{3\alpha^2\sigma_1\left(\|H\|_F^2+\|X\|_F^2\right)}\\ &\le 1.223\gamma\sigma_r\|H\|_F^2 + \gamma\sigma_r\mathrm{tr}^2(H^\top\bar{X})/\sigma_1 + \gamma\sigma_r\|H^\top\bar{X}\|_F^2/\sigma_1.\end{aligned}$$

Note that $\sigma_r \le 1$. Taking $\gamma = 0.166/12.23$ and then combining with (5.4), we obtain

$$\begin{aligned}\langle \nabla f_{\rm ex}(U), H\rangle &\ge 0.1494\sigma_r\|H\|_F^2 + \frac{\sigma_r^3\|\nabla f_{\rm ex}(U)\|_F^2}{222\alpha^2\sigma_1\left(\|H\|_F^2+\|X\|_F^2\right)}\\ &\ge 0.1494\sigma_r\|H\|_F^2 + \frac{\sigma_r^3}{250\alpha^2\sigma_1\|X\|_F^2}\|\nabla f_{\rm ex}(U)\|_F^2,\end{aligned}$$

where we use $\|H\|_F^2 \le \frac{1}{8}\sigma_r \le \frac{1}{8}\|X\|_F^2$ in the last line. Thus we have

$$\langle \nabla f_{\rm ex}(U), H\rangle \ge \frac{1}{\nu}\sigma_r\|H\|_F^2 + \frac{1}{\lambda\|X\|_F^2}\|\nabla f_{\rm ex}(U)\|_F^2$$

for $\nu \ge 7$ and $\lambda \ge 250\alpha^2\sigma_1/\sigma_r^3$ with probability at least $1 - C\exp(-\Omega(n))$, if $m \ge c_0\sigma_r^{-2}nr$ log (c_1r/σ_r). $\square$

6. Numerical Experiments

The purpose of the numerical experiments is the comparison for the exponential-type gradient descent algorithm with the gradient descent algorithm [23]. In our numerical experiments, the target matrix $X \in \mathbb{R}^{n\times r}$ is chosen randomly in standard normal distribution.

Example 6.1. In this example, we test the success rate of the exponential-type gradient descent algorithm with different parameter α. Let $X \in \mathbb{R}^{n\times r}$ with $n = 200, r = 2$, the parameter $\alpha_y = 9$ in spectral initialization and the step size $\mu = 0.1\cdot m/\sum_{i=1}^m y_i$. We test the performance with taking $\alpha = 20$ and 100, respectively. The maximum number of iterations is $T = 3000$. For the measurement number, we vary m within the range $[nr, 4nr]$. For each m, we run 100 times and calculate the success rate. We consider a trial to be successful when the relative error is less than 10^{-5} and the relative error is defined as

$$\min_{O\in\mathcal{O}(r)}\frac{\|XO - U^t\|_F}{\|X\|_F} = \frac{\|XZV^\top - U^t\|_F}{\|X\|_F},$$

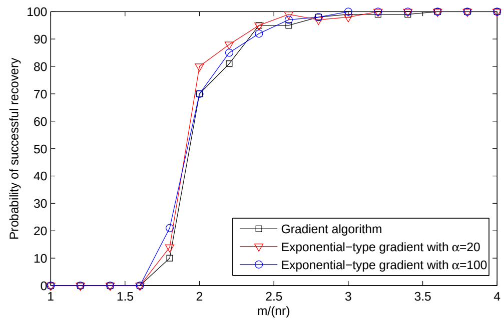
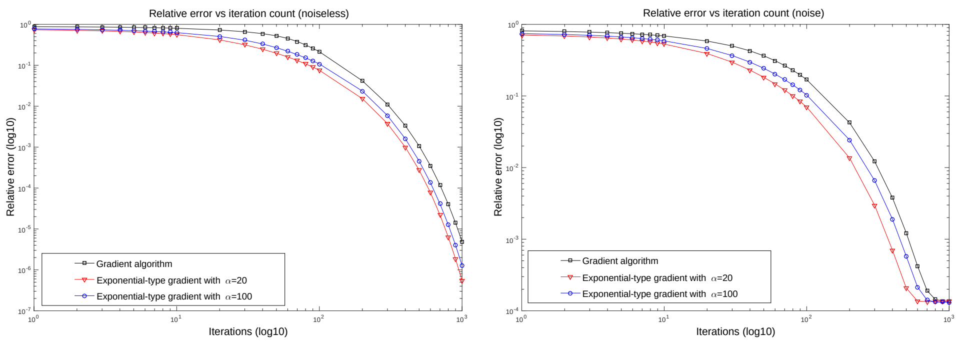


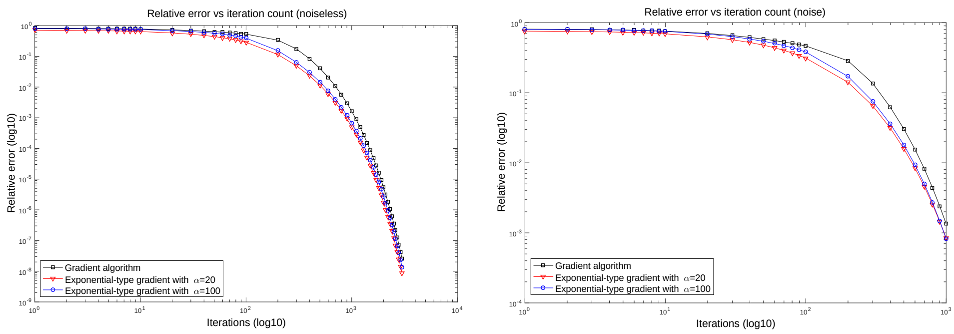
Fig. 6.1.: Success rate experiments: Empirical probability of successful recovery based on 100 random trails for different m/nr. Take $n = 200, r = 2$ and change m/nr between 1 and 4.

where $ZDV^\top$ is the singular value decomposition of $X^\top U^t$. Fig. 6.1 shows the numerical results for exponential-type gradient descent and gradient descent algorithm. The figure shows that exponential-type gradient descent algorithm achieve 100% recovery rate if $m \geq 4nr$ and the empirical success rate is better than the gradient descent algorithm.



(a) (b)

Fig. 6.2.: Convergence experiments: Plot of relative error (log(10)) vs number of iterations (log(10)). Take $n = 200$, $r = 2$ and $m = 3nr$ and the measurement vectors are Gaussian random vectors.



(a) (b)

Fig. 6.3.: Convergence experiments: Plot of relative error (log(10)) vs number of iterations (log(10)). Take $n = 100$, $r = 5$ and $m = 3nr$ and the measurement vectors are Gaussian random vectors.

Example 6.2. In this example, we test the convergence and robustness of the exponential-type gradient descent algorithm. We use noiseless model for (a) to test the convergence and use the noise model for (b) to test the robustness. The noise model is described as $y_i = a_i^\top XX^\top a_i + \epsilon_i$ where the noise $\epsilon_i \sim \mathcal{N}(0, 0.1^2)$, $i = 1, \dots, m$. We take the parameter $\alpha_y = 9$ in spectral initialization and the step size $\mu = 0.1 \cdot m/\sum_{i=1}^m y_i$. Let $X \in \mathbb{R}^{n\times r}$ with $n = 200, r = 2$ (or $n = 100, r = 5$) be generated by standard normal distribution. We take $m = 3nr$. We consider the performance with $\alpha = 20$ and 100, respectively. Figs. 6.2 and 6.3 depict the relative error against the iteration number with different random measurement ensembles. From the figures, we observe that our exponential-type gradient descent algorithm converges faster.

7. Appendix

7.1. Proof of Theorem 4.1

Proof. By homogeneity, it suffices to consider the case where $\|X\|_F = 1$. We assume that $X = (x_1, \dots, x_r) \in \mathbb{R}^{n\times r}$ has orthogonal columns satisfying $\|x_1\|_2 \geq \cdots \geq \|x_r\|_2$. Recall that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are the nonzero eigenvalues of the positive semidefinite matrix $XX^\top$ and then

$$\sigma_j = \|x_j\|_2^2, \quad \text{for } 1 \leq j \leq r.$$

From Lemma 2.3, for $\varepsilon > 0$, we have

$$\frac{1}{m}\sum_{k=1}^m a_k^\top XX^\top a_k = \frac{1}{m}\sum_{k=1}^m y_k \in [1-\varepsilon, 1+\varepsilon], \tag{7.1}$$

with probability at least $1 - 2\exp(-\Omega(n))$, if $m \geq Cn$ where C is a constant depending on ε. Here, we use the fact that $\|XX^\top\|_* = \|X\|_F^2 = 1$. The (7.1) implies that

$$\mathbb{1}_{\{y_i \leq (1-\varepsilon)\alpha_y\}} \leq \mathbb{1}_{\{y_i \leq \frac{\alpha_y}{m}\sum_{k=1}^m y_k\}} \leq \mathbb{1}_{\{y_i \leq (1+\varepsilon)\alpha_y\}}. \tag{7.2}$$

Recall that $Y = \frac{1}{m}\sum_{i=1}^m y_i a_i a_i^\top \mathbb{1}_{\{y_i \leq \frac{\alpha_y}{m}\sum_{k=1}^m y_k\}}$. The (7.2) implies that

$$Y_2 \preceq Y \preceq Y_1 \tag{7.3}$$

holds with high probability where

$$Y_2 := \frac{1}{m}\sum_{i=1}^m y_i a_i a_i^\top \mathbb{1}_{\{y_i \leq (1-\varepsilon)\alpha_y\}}, \quad Y_1 := \frac{1}{m}\sum_{i=1}^m y_i a_i a_i^\top \mathbb{1}_{\{y_i \leq (1+\varepsilon)\alpha_y\}}.$$

We claim the following results:

Claim 7.1. *For any $0 < \delta < 1$, if $\alpha_y \geq C\sqrt{\log(cr\sigma_1/\delta)}$, then*

$$\|\mathbb{E}Y_1 - 2XX^\top - I\|_2 \leq \delta, \qquad \|\mathbb{E}Y_2 - 2XX^\top - I\|_2 \leq \delta. \tag{7.4}$$

The (7.4) implies that $\|\mathbb{E}Y_1\|_2 \geq 1+2\sigma_1-\delta$ and $\|\mathbb{E}Y_2\|_2 \geq 1+2\sigma_1-\delta$. We can use Lemma 2.1 to obtain that if $m \geq C\delta^{-2}(1+2\sigma_1-\delta)^{-2}n$, and then with probability at least $1-4\exp(-\Omega(n))$, we have

$$\|Y_1 - \mathbb{E}Y_1\|_2 \leq \delta, \qquad \|Y_2 - \mathbb{E}Y_2\|_2 \leq \delta, \tag{7.5}$$

where C is a positive constant. Indeed, in Lemma 2.1 we take the i-th row of A as $b_i^\top := \sqrt{y_i}a_i^\top \mathbb{1}_{\{y_i \leq (1+\varepsilon)\alpha_y\}}$ and set $\Sigma = \mathbb{E}Y_1$ with $\|\mathbb{E}Y_1\|_2 \geq 1+2\sigma_1-\delta$ and $t = \delta\|\mathbb{E}Y_1\|_2\sqrt{m}$. Then we can obtain $\|Y_1 - \mathbb{E}Y_1\|_2 \le \delta$. Similarly, we have $\|Y_2 - \mathbb{E}Y_2\|_2 \le \delta$ if we take the i-th row of A as $b_i^\top := \sqrt{y_i} a_i^\top \mathbb{1}_{\{y_i \le (1-\varepsilon)\alpha_y\}}$ and set $\Sigma := \mathbb{E}Y_2$.

Combining (7.3)–(7.5), we have

$$\|Y - 2XX^\top - I\|_2 \le 2\delta \tag{7.6}$$

with probability at least $1 - 6\exp(-\Omega(n))$ provided $m \ge C\delta^{-2}(1 + 2\sigma_1 - \delta)^{-2} n$ and $\alpha_y \ge C\sqrt{\log(cr\sigma_1/\delta)}$. Furthermore, from Wely Theorem we have

$$|\lambda_{r+1} - 1| \le 2\delta \quad \text{and} \quad |\lambda_n - 1| \le 2\delta. \tag{7.7}$$

Next, we turn to consider $d(U_0)$. Recall the definition $U_0 = U\Sigma^{1/2}$ in Algorithm 3.1. Here, $U = (u_1, \ldots, u_r)$ where u_k is normalized eigenvectors corresponding to the eigenvalues λ_k of Y for $k = 1, \ldots, r$, and the scaling of the diagonal matrix Σ is given by $\Sigma_{i,i} = (\lambda_i - \lambda_{r+1})/2$. Hence,

$$\begin{aligned} &\|U_0U_0^\top - XX^\top\|_2 \\ &\le \|U_0U_0^\top - \frac{1}{2}Y + \frac{1}{2}\lambda_{r+1}I\|_2 + \|\frac{1}{2}Y - \frac{1}{2}I - XX^\top\|_2 + \frac{1}{2}\|(\lambda_{r+1} - 1)I\|_2 \\ &\le \frac{1}{2}(\lambda_{r+1} - \lambda_n) + \delta + \frac{1}{2}(\lambda_{r+1} - 1) \le 4\delta, \end{aligned}$$

where the second inequality follows from (7.6) and the last inequality follows from (7.7). Then, using the following fact (see, e.g. the Initialization of [27])

$$\min_{O \in \mathcal{O}(r)} \|U_0 - XO\|_F^2 \le \frac{\|U_0U_0^\top - XX^\top\|_F^2}{(2\sqrt{2} - 2)\sigma_r},$$

and taking $\delta \le \frac{\sigma_r}{18\sqrt{r}}$, we obtain

$$\min_{O \in \mathcal{O}(r)} \|U_0 - XO\|_F^2 \le \frac{2r\|U_0U_0^\top - XX^\top\|_2^2}{(2\sqrt{2} - 2)\sigma_r} \le \frac{32r\delta^2}{(2\sqrt{2} - 2)\sigma_r} \le \frac{\sigma_r}{8},$$

where we use $\|A\|_F \le \sqrt{\operatorname{rank}(A)}\|A\|_2$ in the first inequality. The choice of δ implies that the measurements $m \ge C\sigma_r^{-2} nr$ and $\alpha_y \ge C\sqrt{\log(c'\kappa r)}$, where $\kappa = \sigma_1/\sigma_r$ denotes the ratio of the largest to the smallest nonzero eigenvalues of matrix $XX^\top$.

We remain to prove Claim 7.1. There exists an orthogonal matrix $O \in \mathbb{R}^{r \times r}$ such that $X = O(\|x_1\|_2 e_1, \ldots, \|x_r\|_2 e_r)$. Then

$$\begin{aligned} O^\top(\mathbb{E}Y_1 - 2XX^\top - I)O &= O^\top \mathbb{E}Y_1 O - \left(2\sum_{k=1}^r \|x_k\|_2^2 e_k e_k^\top + I\right), \\ O^\top \mathbb{E}Y_1 O &= \mathbb{E}\left[\sum_{k=1}^r \|x_k\|_2^2 a_{i,k}^2 a_i a_i^\top \mathbb{1}_{\{\sum_{k=1}^r \|x_k\|_2^2 a_{i,k}^2 \le (1+\varepsilon)\alpha_y\}}\right]. \end{aligned} \tag{7.8}$$

A simple calculation is that

$$\mathbb{E}\left[\sum_{k=1}^r \|x_k\|_2^2 a_{i,k}^2 a_i a_i^\top\right] = 2\sum_{k=1}^r \|x_k\|_2^2 e_k e_k^\top + I, \tag{7.9}$$

which implies that

$$O^\top \mathbb{E}Y_1 O \le 2\sum_{k=1}^r \|x_k\|_2^2 e_k e_k^\top + I, \tag{7.10}$$

where we write $M_2 \le M_1$ if all entries of $M_1 - M_2$ are nonnegative. On the other hand, from (7.8) we obtain that

$$O^\top \mathbb{E} Y_1 O = \mathbb{E}\left[\sum_{k=1}^r \|x_k\|_2^2 a_{i,k}^2 a_i a_i^\top\right] - \mathbb{E}\left[\sum_{k=1}^r \|x_k\|_2^2 a_{i,k}^2 a_i a_i^\top \mathbb{1}_{\{\sum_{k=1}^r \|x_k\|_2^2 a_{i,k}^2 \ge (1+\varepsilon)\alpha_y\}}\right]. \quad (7.11)$$

For any $1 \le j, l, k \le r$ and $\delta > 0$, by Hölder's inequality we have

$$\begin{aligned}
&\mathbb{E}\left[\|x_k\|_2^2 a_{i,j}^2 a_{i,l}^2 \mathbb{1}_{\{\sum_{k=1}^r \|x_k\|_2^2 a_{i,k}^2 \ge (1+\varepsilon)\alpha_y\}}\right] \\
\le & \|x_1\|_2^2 \sqrt{\mathbb{E}[a_{i,j}^4 a_{i,l}^4]} \cdot \sqrt{\mathbb{P}\left\{\sum_{k=1}^r \|x_k\|_2^2 a_{i,k}^2 \ge (1+\varepsilon)\alpha_y\right\}} \\
\le & C_1 \|x_1\|_2^2 \exp\left(-C_0 \min\left(\frac{(1+\varepsilon)^2\alpha_y^2}{\|x_1\|_2^4 + \cdots + \|x_r\|_2^4}, \frac{(1+\varepsilon)\alpha_y}{\|x_1\|_2^2}\right)\right) \\
\le & C_1 \sigma_1 \exp\left(-C_0 (1+\varepsilon)^2 \alpha_y^2\right) \le \frac{\delta}{r} \qquad (7.12)
\end{aligned}$$

provided $\alpha_y \ge C\sqrt{\log(cr\sigma_1/\delta)}$, where the second inequality follows from Lemma 2.2 and the third inequality follows from the fact that $\|X\|_F = 1$ and $\|x_r\|_2 \le \cdots \le \|x_1\|_2 \le 1$. The (7.12) implies that

$$\mathbb{E}\left[\sum_{k=1}^r \|x_k\|_2^2 a_{i,k}^2 a_i a_i^\top \mathbb{1}_{\{\sum_{k=1}^r \|x_k\|_2^2 a_{i,k}^2 \ge (1+\varepsilon)\alpha_y\}}\right] \le \delta I. \quad (7.13)$$

Thus, combining (7.9), (7.11) and (7.13) we have

$$O^\top \mathbb{E} Y_1 O \ \ge 2 \sum_{k=1}^r \|x_k\|_2^2 e_k e_k^\top + (1-\delta) I. \quad (7.14)$$

Combining (7.10) and (7.14) and noting that $O^\top \mathbb{E} Y_1 O$ is a diagonal matrix, we obtain

$$\|\mathbb{E} Y_1 - 2XX^\top - I\|_2 = \|O^\top(\mathbb{E} Y_1 - 2XX^\top - I)O\|_2 \le \delta.$$

Similarly, we can obtain $\|\mathbb{E} Y_2 - 2XX^\top - I\|_2 \le \delta$, which completes the proof. □

7.2. Proof of Proposition 5.1

We always assume that $\|X\|_F = 1$ throughout the proof. We set $H := U - \bar{X}$ where $\bar{X} = \underset{Z \in \mathcal{X}}{\operatorname{argmin}} \|U - Z\|_F$ and $\mathcal{X}$ is the solution set. Then the exponential-type gradient descent algorithm can be rewritten as

$$\nabla f_{\mathrm{ex}}(U) = \frac{1}{m}\sum_{i=1}^m (a_i^\top H H^\top a_i + 2a_i^\top H \bar{X}^\top a_i)(a_i a_i^\top H + a_i a_i^\top \bar{X}) \cdot \exp\left(-\frac{m y_i}{\alpha \sum_{k=1}^m y_k}\right). \quad (7.15)$$

For convenience, we let

$$\rho_{i,\alpha} := \exp\left(-\frac{m y_i}{\alpha \sum_{i=1}^m y_i}\right), \ i = 1, \ldots, m. \quad (7.16)$$

To prove Proposition 5.1, we need the following lemmas.

Lemma 7.1. *For any fixed* $\alpha \geq 20$ *and* $\delta > 0$*, if* $m \geq c_0\alpha^2\delta^{-2}nr\log(\sqrt{r}/\delta)$*, then with probability at least* $1 - C\exp(-\Omega(\alpha^{-2}\delta^2 m))$*, the followings hold for all non-zero matrix* $U \in \mathbb{R}^{n\times r}$*:*

$$(a)\ \frac{1}{m}\sum_{i=1}^{m}(a_i^\top H\bar{X}^\top a_i)^2\rho_{i,\alpha} \geq (0.78\sigma_r - 2\delta)\|H\|_F^2 + 0.78\mathrm{tr}^2(H^\top\bar{X}) + 0.78\|H^\top\bar{X}\|_F^2,$$

$$(b)\ \frac{1}{m}\sum_{i=1}^{m}(a_i^\top H\bar{X}^\top a_i)^2\rho_{i,\alpha} \leq (\sigma_1 + 2\delta)\|H\|_F^2 + \mathrm{tr}^2(H^\top\bar{X}) + \|H^\top\bar{X}\|_F^2,$$

where C, c_0 *are universal constants.*

Proof. Suppose for the moment that H is independent from a_i. By homogeneity, it suffices to establish the claim for the case $\|H\|_F = 1$. From (7.1) we have

$$\exp\Big(-\frac{a_i^\top XX^\top a_i}{0.99\alpha}\Big) \leq \rho_{i,\alpha} \leq \exp\Big(-\frac{a_i^\top XX^\top a_i}{1.01\alpha}\Big) \tag{7.17}$$

with high probability. For convenience, we set

$$\bar{\rho}_{i,\alpha} := \exp\Big(-\frac{a_i^\top \bar{X}\bar{X}^\top a_i}{0.99\alpha}\Big), \qquad i = 1,\ldots,m. \tag{7.18}$$

Noting that $a_i^\top\bar{X}\bar{X}^\top a_i = a_i^\top XX^\top a_i$, we have

$$\frac{1}{m}\sum_{i=1}^{m}(a_i^\top H\bar{X}^\top a_i)^2\rho_{i,\alpha} \geq \frac{1}{m}\sum_{i=1}^{m}(a_i^\top H\bar{X}^\top a_i)^2\bar{\rho}_{i,\alpha}. \tag{7.19}$$

We claim the following results:

Claim 7.2. *For any fixed parameter* $\alpha \geq 20$ *it holds*

1) $\mathbb{E}\left[(a_i^\top H\bar{X}^\top a_i)^2\right] \geq \sigma_r\|H\|_F^2 + \mathrm{tr}^2(H^\top\bar{X}) + \|H^\top\bar{X}\|_F^2$

2) $\mathbb{E}\left[(a_i^\top H\bar{X}^\top a_i)^2\right] \leq \sigma_1\|H\|_F^2 + \mathrm{tr}^2(H^\top\bar{X}) + \|H^\top\bar{X}\|_F^2$

3) $\mathbb{E}\left[(a_i^\top H\bar{X}^\top a_i)^2\bar{\rho}_{i,\alpha}\right] \geq 0.78\mathbb{E}\left[(a_i^\top H\bar{X}^\top a_i)^2\right]$.

Then combining 3) and 1) we obtain that

$$\mathbb{E}\left[(a_i^\top H\bar{X}^\top a_i)^2\bar{\rho}_{i,\alpha}\right] \geq 0.78\sigma_r\|H\|_F^2 + 0.78\mathrm{tr}^2(H^\top\bar{X}) + 0.78\|H^\top\bar{X}\|_F^2.$$

Since

$$(a_i^\top H\bar{X}^\top a_i)^2\bar{\rho}_{i,\alpha} \leq (a_i^\top\bar{X}\bar{X}^\top a_i)\bar{\rho}_{i,\alpha}(a_i^\top HH^\top a_i)$$

and $(a_i^\top\bar{X}\bar{X}^\top a_i)\bar{\rho}_{i,\alpha}$ is bounded, it means that $(a_i^\top H\bar{X}^\top a_i)^2\bar{\rho}_{i,\alpha}$ is a sub-exponential random variable with ψ_1 norm $O(\alpha\|H\|_F^2)$. We can use Lemma 2.2 to obtain that

$$\begin{aligned}\frac{1}{m}\sum_{i=1}^{m}(a_i^\top H\bar{X}^\top a_i)^2\bar{\rho}_{i,\alpha} &\geq \mathbb{E}\big[(a_i^\top H\bar{X}^\top a_i)^2\bar{\rho}_{i,\alpha}\big] - \delta\|H\|_F^2\\ &\geq (0.78\sigma_r - \delta)\|H\|_F^2 + 0.78\mathrm{tr}^2(H^\top\bar{X}) + 0.78\|H^\top\bar{X}\|_F^2 \end{aligned} \tag{7.20}$$

holds with probability at least $1-\exp(-\Omega(\alpha^{-2}\delta^2 m))$ where $\delta>0$. Combining (7.19) and (7.20), we obtain that (a) holds for a fixed $H\in\mathbb{R}^{n\times r}$.

We construct an ϵ-net $\mathcal{N}_\epsilon\subset\mathbb{R}^{n\times r}$ with cardinality $|\mathcal{N}_\epsilon|\le(1+\frac{2}{\epsilon})^{nr}$ such that for any $H\in\mathbb{R}^{n\times r}$ with $\|H\|_F=1$, there exists $H_0\in\mathcal{N}_\epsilon$ satisfying $\|H-H_0\|_F\le\epsilon$. Taking a union bound over this set gives that

$$\frac{1}{m}\sum_{i=1}^m(a_i^\top H_0\bar{X}^\top a_i)^2\bar{\rho}_{i,\alpha}\ge(0.78\sigma_r-\delta)\|H_0\|_F^2+0.78\mathrm{tr}^2(H_0^\top\bar{X})+0.78\|H_0^\top\bar{X}\|_F^2$$

holds for all $H_0\in\mathcal{N}_\epsilon$ with probability at least $1-(1+\frac{2}{\epsilon})^{nr}\exp(-\Omega(\alpha^{-2}\delta^2 m))$. Note that $\bar{\rho}_{i,\alpha}<1$ for all i. Then there exists a universal constant $c_1>0$ such that

$$\begin{aligned}&\left|\frac{1}{m}\sum_{i=1}^m(a_i^\top H\bar{X}^\top a_i)^2\bar{\rho}_{i,\alpha}-\frac{1}{m}\sum_{i=1}^m(a_i^\top H_0\bar{X}^\top a_i)^2\bar{\rho}_{i,\alpha}\right|\\ &\le\frac{1}{m}\sum_{i=1}^m\left|a_i^\top H\bar{X}^\top a_i-a_i^\top H_0\bar{X}^\top a_i\right|\\ &\le c_1\|HX^\top-H_0X^\top\|_*\le c_1\sqrt{r}\|H-H_0\|_F\le c_1\sqrt{r}\epsilon,\end{aligned}\tag{7.21}$$

where we use Lemma 2.3 in the second line, the fact $\|A\|_*\le\sqrt{\mathrm{rank}(A)}\|A\|_F$ in the third line. Indeed, according to Lemma 2.3, for any $\delta\in(0,1)$, if $m\ge c_0\delta^{-2}n$, then with probability at least $1-C\exp(-\Omega(n))$ we have

$$\begin{aligned}&\frac{1}{m}\sum_{i=1}^m\left|a_i^\top HX^\top a_i-a_i^\top H_0X^\top a_i\right|\\ &\le(1+\delta)\|HX^\top-H_0X^\top\|_*\le c_1\|HX^\top-H_0X^\top\|_*.\end{aligned}$$

By choosing $\epsilon=\frac{\delta}{c_1\sqrt{r}}$ in (7.21), we conclude the first part of lemma.

We now turn to the part (b). The estimate (7.17) implies that

$$\rho_{i,\alpha}\le\exp\left(-\frac{a_i^\top XX^\top a_i}{1.01\alpha}\right)$$

holds with high probability. It gives that

$$\frac{1}{m}\sum_{i=1}^m(a_i^\top H\bar{X}^\top a_i)^2\rho_{i,\alpha}\le\frac{1}{m}\sum_{i=1}^m(a_i^\top H\bar{X}^\top a_i)^2\exp\left(-\frac{a_i^\top XX^\top a_i}{1.01\alpha}\right).$$

From Claim 7.2, we have

$$\mathbb{E}\left[(a_i^\top H\bar{X}^\top a_i)^2\exp\left(-\frac{a_i^\top XX^\top a_i}{1.01\alpha}\right)\right]\le\sigma_1\|H\|_F^2+\mathrm{tr}^2(H^\top\bar{X})+\|H^\top\bar{X}\|_F^2.$$

Similarly, $(a_i^\top H\bar{X}^\top a_i)^2\exp\left(-\frac{a_i^\top XX^\top a_i}{1.01\alpha}\right)$ is a sub-exponential random variable with sub-exponential norm $O(\alpha\|H\|_F^2)$. Then, we can employ the method for proving part (a) to prove part (b).

Lemma 7.2. *For a fixed $\lambda>0$, for any $H\in\mathbb{R}^{n\times r}$ and $\delta>0$, if $m\ge c_0\delta^{-2}\lambda^{-2}nr\log(\sqrt{r}/(\delta\lambda))$, then with probability at least $1-C\exp(-\Omega(\delta^2\lambda^2 m))$, we have*

$$\frac{1}{m}\sum_{i=1}^m(a_i^\top HH^\top a_i)^2\exp\left(-\lambda\frac{a_i^\top HH^\top a_i}{\|H\|_F^2}\right)\le2\|HH^\top\|_F^2+(2\delta+1)\|H\|_F^4.$$

Here, c_0, C are some universal constants.

Proof. Without loss of generality, we only need to prove the lemma in the case $\|H\|_F = 1$. It is straightforward to show that

$$\mathbb{E}\left[(a_i^\top HH^\top a_i)^2 \exp\left(-\lambda a_i^\top HH^\top a_i\right)\right] \le \mathbb{E}\left[(a_i^\top HH^\top a_i)^2\right] = 2\|HH^\top\|_F^2 + \|H\|_F^4.$$

Observe that $(a_i^\top HH^\top a_i)^2 \exp\left(-\lambda a_i^\top HH^\top a_i\right)$ is a sub-exponential random variable with sub-exponential norm $O(1/\lambda \cdot \|H\|_F^2)$. According to Lemma 2.2 we have

$$\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)^2 \exp\left(-\lambda a_i^\top HH^\top a_i\right) \le 2\|HH^\top\|_F^2 + \|H\|_F^4 + \frac{\delta_0}{\lambda}\|H\|_F^2$$

with probability $1-\exp(-\Omega(\delta_0^2 m))$. We next construct an ϵ-net $\mathcal{N}_\epsilon$ with $|\mathcal{N}_\epsilon| \le (1+\frac{2}{\epsilon})^{nr}$ such that for any $H \in \mathbb{R}^{n\times r}$ with $\|H\|_F = 1$, there exists $H_0 \in \mathcal{N}_\epsilon$ satisfying $\|H - H_0\|_F \le \epsilon$. Since $x^2 e^{-\lambda x}$ is Lipschitz function with Lipschitz constant $O(1/\lambda^2)$, we have

$$\begin{aligned}
&\Big|\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)^2 \exp\Big(-\lambda a_i^\top HH^\top a_i\Big) - \frac{1}{m}\sum_{i=1}^m (a_i^\top H_0 H_0^\top a_i)^2 \exp\Big(-\lambda a_i^\top H_0 H_0^\top a_i\Big)\Big| \\
\le& \frac{1}{\lambda^2 m}\sum_{i=1}^m \Big|a_i^\top HH^\top a_i - a_i^\top H_0 H_0^\top a_i\Big| \le \frac{c_2\sqrt{r}\epsilon}{\lambda^2},
\end{aligned}$$

where the last inequality follows from Lemma 2.3. By choosing $\epsilon = \frac{\delta_0\lambda}{c_2\sqrt{r}}$, we obtain

$$\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)^2 \exp\left(-\lambda a_i^\top HH^\top a_i\right) \le 2\|HH^\top\|_F^2 + \|H\|_F^4 + \frac{2\delta_0}{\lambda}\|H\|_F^2$$

with probability at least $1-\exp(-\Omega(\delta_0^2 m))$ if $m \ge c_0\delta_0^{-2} nr\log(\sqrt{r}/(\delta_0\lambda))$. Finally, noting that $\|H\|_F = 1$ and taking $\delta_0 = \lambda\delta$, we arrive at the conclusion. □

Corollary 7.1. *For any* $\delta > 0$, $U \in \mathbb{R}^{n\times r}$ *and* $H = U - \bar{X}$, *if* $m \ge c_0\alpha^2\delta^{-2}\sigma_r^{-2} nr\log(\alpha\sqrt{r}/(\delta\sigma_r))$, *then with probability at least* $1 - C\exp(-\Omega(n))$, *it holds*

$$\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)^2 \rho_{i,\alpha} \le 2\|HH^\top\|_F^2 + (2\delta+1)\|H\|_F^4.$$

Proof. Since σ_r is the smallest eigenvalue of $XX^\top$, we have

$$y_i = a_i^\top XX^\top a_i \ge \sigma_r\|a_i\|^2,$$

which implies that

$$\|a_i\|^2 \le \frac{a_i^\top XX^\top a_i}{\sigma_r} = \frac{y_i}{\sigma_r}. \tag{7.22}$$

On the other hand, we have

$$a_i^\top HH^\top a_i \le \|H\|_F^2\|a_i\|^2. \tag{7.23}$$

Combining (7.22) and (7.23), we obtain that

$$y_i \ge \sigma_r\frac{a_i^\top HH^\top a_i}{\|H\|_F^2}. \tag{7.24}$$

According to (7.17) and (7.24), we obtain that

$$\frac{1}{m}\sum_{i=1}^{m}(a_i^\top HH^\top a_i)^2\rho_{i,\alpha} \le \frac{1}{m}\sum_{i=1}^{m}(a_i^\top HH^\top a_i)^2 \exp\left(-\frac{\sigma_r}{1.01\alpha}\cdot\frac{a_i^\top HH^\top a_i}{\|H\|_F^2}\right).$$

We take $\lambda = \frac{\sigma_r}{1.01\alpha}$ in Lemma 7.2 and arrive at the conclusion. □

Proof of Proposition 5.1. To state conveniently, we set

$$\beta^2 = \frac{1}{m}\sum_{i=1}^{m}(a_i^\top HH^\top a_i)^2\rho_{i,\alpha}, \quad \gamma^2 = \frac{2}{m}\sum_{i=1}^{m}(a_i^\top H\bar{X}^\top a_i)^2\rho_{i,\alpha}.$$

According to the expression of exponential-type gradient (7.15), we have

$$\begin{aligned}
&\langle \nabla f_{\mathrm{ex}}(U), H\rangle = \beta^2+\gamma^2+\frac{3}{m}\sum_{i=1}^{m}(a_i^\top H\bar{X}^\top a_i)(a_i^\top HH^\top a_i)\rho_{i,\alpha}\\
\ge&\beta^2+\gamma^2-\frac{3}{m}\sqrt{\sum_{i=1}^{m}(a_i^\top H\bar{X}^\top a_i)^2\rho_{i,\alpha}}\cdot\sqrt{\sum_{i=1}^{m}(a_i^\top HH^\top a_i)^2\rho_{i,\alpha}}\\
=&\beta^2+\gamma^2-\frac{3}{\sqrt{2}}\beta\gamma = \Big(\gamma-\frac{3}{2\sqrt{2}}\beta\Big)^2-\frac{1}{8}\beta^2\\
\ge&\Big(\frac{\gamma^2}{2}-\frac{9}{8}\beta^2\Big)-\frac{1}{8}\beta^2 = \frac{\gamma^2}{2}-\frac{5}{4}\beta^2\\
=&\frac{1}{m}\sum_{i=1}^{m}(a_i^\top H\bar{X}^\top a_i)^2\rho_{i,\alpha}-\frac{5}{4m}\sum_{i=1}^{m}(a_i^\top HH^\top a_i)^2\rho_{i,\alpha}\\
\ge&(0.78\sigma_r-2\delta_1)\|H\|_F^2+0.78\mathrm{tr}^2(H^\top\bar{X})+0.78\|H^\top\bar{X}\|_F^2-\frac{5}{2}\|HH^\top\|_F^2-\frac{5(2\delta_2+1)}{4}\|H\|_F^4\\
\ge&\left(0.78\sigma_r-2\delta_1-\frac{5(2\delta_2+3)}{4}\|H\|_F^2\right)\|H\|_F^2+0.78\left(\mathrm{tr}^2(H^\top\bar{X})+\|H^\top\bar{X}\|_F^2\right),
\end{aligned}$$

where we use Cauchy-Schwarz inequality in the second line, the inequality $(\gamma-\beta)^2 \ge \frac{\gamma^2}{2}-\beta^2$ in the fourth line, Lemma 7.1 and Corollary 7.1 in the sixth line, and the fact that $\|HH^\top\|_F \le \|H\|_F^2$ in the last line. Note that $\|H\|_F^2 = \|U-\bar{X}\|_F^2 = d(U)^2 \le \frac{1}{8}\sigma_r$. Taking $\delta_1 \le \frac{1}{16}\sigma_r$ and $\delta_2 \le \frac{1}{16}$, we obtain that

$$\langle \nabla f_{\mathrm{ex}}(U), H\rangle \ge 0.166\sigma_r\|H\|_F^2+0.78\left(\mathrm{tr}^2(H^\top\bar{X})+\|H^\top\bar{X}\|_F^2\right)$$

with probability at least $1-C\exp(-\Omega(n)$, if $m \ge c_0\sigma_r^{-2}nr\log(c_1 r/\sigma_r)$. This implies the part (a) holds. Next, we turn to the part (b). We consider

$$\|\nabla f_{\mathrm{ex}}(U)\|_F^2 = \max_{\|W\|_F=1, W\in\mathbb{R}^{n\times r}}|\langle\nabla f_{\mathrm{ex}}(U), W\rangle|^2$$

on the case where $H = U - \bar{X} \leq \sqrt{\frac{1}{8}\sigma_r}$. Recall the notation $\rho_{i,\alpha}$ in formula (7.16), and we have

$$\begin{aligned}
&|\langle \nabla f_{\mathrm{ex}}(U), W\rangle|^2 \\
=&\left(\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)(a_i^\top HW^\top a_i)\rho_{i,\alpha} + \frac{2}{m}\sum_{i=1}^m (a_i^\top H\bar{X}^\top a_i)(a_i^\top HW^\top a_i)\rho_{i,\alpha}\right. \\
&\left.+\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)(a_i^\top \bar{X}W^\top a_i)\rho_{i,\alpha} + \frac{2}{m}\sum_{i=1}^m (a_i^\top H\bar{X}^\top a_i)(a_i^\top \bar{X}W^\top a_i)\rho_{i,\alpha}\right)^2 \\
\leq& 4\left(\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)(a_i^\top HW^\top a_i)\rho_{i,\alpha}\right)^2 + 16\left(\frac{1}{m}\sum_{i=1}^m (a_i^\top H\bar{X}^\top a_i)(a_i^\top HW^\top a_i)\rho_{i,\alpha}\right)^2 \\
&+ 4\left(\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)(a_i^\top \bar{X}W^\top a_i)\rho_{i,\alpha}\right)^2 + 16\left(\frac{1}{m}\sum_{i=1}^m (a_i^\top H\bar{X}^\top a_i)(a_i^\top \bar{X}W^\top a_i)\rho_{i,\alpha}\right)^2.
\end{aligned}$$

We first consider the term $4\left(\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)(a_i^\top HW^\top a_i)\rho_{i,\alpha}\right)^2$. Using Cauchy-Schwarz inequality, we obtain that

$$\begin{aligned}
&4\left(\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)(a_i^\top HW^\top a_i)\rho_{i,\alpha}\right)^2 \\
\leq& 4\left(\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)^2\rho_{i,\alpha}\right)\left(\frac{1}{m}\sum_{i=1}^m (a_i^\top HW^\top a_i)^2\rho_{i,\alpha}\right) \\
\leq& 4\left(\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)^2\rho_{i,\alpha}\right)\left(\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)(a_i^\top WW^\top a_i)\rho_{i,\alpha}\right).
\end{aligned}$$

According to Corollary 7.1, we have

$$\frac{1}{m}\sum_{i=1}^m (a_i^\top HH^\top a_i)^2\rho_{i,\alpha} \leq \left(2\|HH^\top\|_F^2 + (2\delta_2+1)\|H\|_F^4\right) \tag{7.25}$$

with probability at least $1 - C\exp(-\Omega(n))$ provided $m \geq c_0\delta_2^{-2}\sigma_r^{-2}nr\log(\sqrt{r}/(\delta_2\sigma_r))$. Noting that $a_i^\top XX^\top a_i \geq \sigma_r\|a_i\|^2$ and $a_i^\top HH^\top a_i \leq \|H\|_F^2\|a_i\|^2$ we have

$$\frac{a_i^\top XX^\top a_i}{2.02\alpha} \geq \frac{\sigma_r \cdot a_i^\top HH^\top a_i}{2.02\alpha\|H\|_F^2} \quad \text{and} \quad \frac{a_i^\top XX^\top a_i}{2.02\alpha} \geq \frac{\sigma_r \cdot a_i^\top WW^\top a_i}{2.02\alpha}.$$

It gives that

$$\begin{aligned}
&(a_i^\top HH^\top a_i)(a_i^\top WW^\top a_i)\rho_{i,\alpha} \\
\leq&(a_i^\top HH^\top a_i)(a_i^\top WW^\top a_i)\exp\left(-\frac{a_i^\top XX^\top a_i}{1.01\alpha}\right) \\
\leq&(a_i^\top HH^\top a_i)\exp\left(-\frac{\sigma_r \cdot a_i^\top HH^\top a_i}{2.02\alpha\|H\|_F^2}\right)(a_i^\top WW^\top a_i)\exp\left(-\frac{\sigma_r \cdot a_i^\top WW^\top a_i}{2.02\alpha}\right) \\
\leq&\|H\|_F^2\left(\frac{1.01\alpha}{e\sigma_r}\right)^2
\end{aligned} \tag{7.26}$$

where we use inequality $xe^{-\gamma x} \leq 1/(e\gamma)$ for any $x \geq 0$ in the last line. Combining formulas (7.25) and (7.26), we obtain

$$\begin{aligned}
&4\left(\frac{1}{m}\sum_{i=1}^{m}(a_i^\top HH^\top a_i)(a_i^\top HW^\top a_i)\rho_{i,\alpha}\right)^2\\
\leq &4\Big(\frac{1.01\alpha}{e\sigma_r}\Big)^2\|H\|_F^2\left(2\|HH^\top\|_F^2+(2\delta_2+1)\|H\|_F^4\right).
\end{aligned}$$

The other three terms can be bounded similarly. For the second term, we have

$$\begin{aligned}
&16\left(\frac{1}{m}\sum_{i=1}^{m}(a_i^\top H\bar{X}^\top a_i)(a_i^\top HW^\top a_i)\rho_{i,\alpha}\right)^2\\
\leq &16\left(\frac{1}{m}\sum_{i=1}^{m}(a_i^\top H\bar{X}^\top a_i)^2\rho_{i,\alpha}\right)\left(\frac{1}{m}\sum_{i=1}^{m}(a_i^\top HW^\top a_i)^2\rho_{i,\alpha}\right)\\
\leq &4\Big(\frac{1.01\alpha}{e\sigma_r}\Big)^2\|H\|_F^2\left(4(\sigma_1+2\delta_1)\|H\|_F^2+4\mathrm{tr}^2(H^\top\bar{X})+4\|H^\top\bar{X}\|_F^2\right)
\end{aligned}$$

with probability at least $1-C\exp(-\Omega(n))$ provided $m \geq c_0\delta_1^{-2}nr\log(\sqrt{r}/\delta_1)$, where we use the part (b) of Lemma 7.1 in the last line. The third term and fourth term can be bounded as

$$\begin{aligned}
&4\left(\frac{1}{m}\sum_{i=1}^{m}(a_i^\top HH^\top a_i)(a_i^\top \bar{X}W^\top a_i)\rho_{i,\alpha}\right)^2\\
&\quad\leq 4\Big(\frac{1.01\alpha}{e\sigma_r}\Big)^2\|X\|_F^2\left(2\|HH^\top\|_F^2+(2\delta_2+1)\|H\|_F^4\right),\\
&\left(\frac{1}{m}\sum_{i=1}^{m}(a_i^\top H\bar{X}^\top a_i)(a_i^\top \bar{X}W^\top a_i)\rho_{i,\alpha}\right)^2\\
&\quad\leq \Big(\frac{1.01\alpha}{2e\sigma_r}\Big)^2\|X\|_F^2\left(4(\sigma_1+2\delta_1)\|H\|_F^2+4\mathrm{tr}^2(H^\top\bar{X})+4\|H^\top\bar{X}\|_F^2\right).
\end{aligned}$$

Putting there inequalities together and noting that $\|HH^\top\|_F \leq \|H\|_F^2$, we have

$$\begin{aligned}
\|\nabla f_{\mathrm{ex}}(U)\|_F^2 \leq \Big(\frac{2.02\alpha}{e\sigma_r}\Big)^2&\left(\|H\|_F^2+\|X\|_F^2\right)\Big(\left(4\sigma_1+8\delta_1+(2\delta_2+3)\|H\|_F^2\right)\|H\|_F^2\\
&+4\mathrm{tr}^2(H^\top\bar{X})+4\|H^\top\bar{X}\|_F^2\Big).
\end{aligned}$$

Furthermore, noticing that $\|H\|_F^2 \leq \frac{1}{8}\sigma_r$ and choosing $\delta_1 \leq \frac{1}{16}\sigma_r$, $\delta_2 \leq \frac{1}{16}$, it follows that

$$\frac{\sigma_r^2\|\nabla f_{\mathrm{ex}}(U)\|_F^2}{3\alpha^2\left(\|H\|_F^2+\|X\|_F^2\right)} \leq 1.223\sigma_1\|H\|_F^2+\mathrm{tr}^2(H^\top\bar{X})+\|H^\top\bar{X}\|_F^2$$

with probability at least $1-C\exp(-\Omega(n)$, if $m \geq c_0\sigma_r^{-2}nr\log(c_1r/\sigma_r)$.

The rest paper is to check the Claim 7.2. For 1) and 2) of the Claim 7.2, let $O_1 = argmin_{O\in\mathcal{O}(r)}\|U-XO\|_F$, then $\bar{X} = XO_1$. Recall that X has orthogonal column vectors, and then there exists an orthogonal matrix $O_2 \in \mathbb{R}^{n\times n}$ such that $X = O_2(\|x_1\|e_1,\ldots,\|x_r\|e_r)$.

Let $\hat{H} := HO_1^\top$, $\tilde{H} = O_2^\top \hat{H}$ and $\hat{h}_s, \tilde{h}_s, x_s$ denote the sth column of $\hat{H}, \tilde{H}, X$ respectively, and $a_{i,s}$ denotes the sth entry of a_i. It follows that

$$\begin{aligned}
&\mathbb{E}\left[(a_i^\top H\bar{X}^\top a_i)^2\right] = \mathbb{E}\left[(a_i^\top \hat{H}X^\top a_i)^2\right] = \mathbb{E}(a_i^\top O_2\tilde{H}X^\top O_2O_2^\top a_i)\\
&= \mathbb{E}(a_i^\top \tilde{H}X^\top O_2 a_i) = \mathbb{E}\left[\|x_1\|(\tilde{h}_1^\top a_i)a_{i,1} + \cdots + \|x_r\|(\tilde{h}_r^\top a_i)a_{i,r}\right]^2\\
&= \mathbb{E}\left[\sum_{s=1}^r \|x_s\|^2(\tilde{h}_s^\top a_i)^2 a_{i,s}^2 + \sum_{s\neq k}\|x_s\|\|x_k\|(\tilde{h}_s^\top a_i)(\tilde{h}_k^\top a_i)a_{i,s}a_{i,k}\right]\\
&= \sum_{s=1}^r \left(\|x_s\|^2\|\tilde{h}_s\|^2 + 2\|x_s\|^2\tilde{h}_{s,s}^2\right) + \sum_{s\neq k}\|x_s\|\|x_k\|\left(\tilde{h}_{s,s}\tilde{h}_{k,k} + \tilde{h}_{s,k}\tilde{h}_{k,s}\right) \qquad (7.27)\\
&= \sum_{s=1}^r \|x_s\|^2\|\hat{h}_s\|^2 + \sum_{s,k}\|x_s\|\|x_k\|\left(\tilde{h}_s^\top e_s\tilde{h}_k^\top e_k + \tilde{h}_s^\top e_k\tilde{h}_k^\top e_s\right)\\
&= \sum_{s=1}^r \|x_s\|^2\|\hat{h}_s\|^2 + \sum_{s,k}(x_s^\top \hat{h}_s x_k^\top \hat{h}_k + x_s^\top \hat{h}_k x_k^\top \hat{h}_s)\\
&\geq \sigma_r\|\hat{H}\|_F^2 + \mathrm{tr}^2(X^\top \hat{H}) + \mathrm{tr}(X^\top \hat{H}X^\top \hat{H})\\
&= \sigma_r\|H\|_F^2 + \mathrm{tr}^2(H^\top \bar{X}) + \mathrm{tr}(H^\top \bar{X}H^\top \bar{X})\\
&= \sigma_r\|H\|_F^2 + \mathrm{tr}^2(H^\top \bar{X}) + \|H^\top \bar{X}\|_F^2, \qquad (7.28)
\end{aligned}$$

where the last equation follows from that $H^\top \bar{X}$ is a symmetric matrix and the symmetry of $HX^\top = (U-\bar{X})X^\top$ can be seen by the singular-value decomposition of $X^\top U$. More specifically, suppose that the singular-value decomposition of $X^\top U$ is $WDV^\top$, then we have

$$O_1 := \operatorname*{argmin}_{O\in\mathcal{O}(r)} \|U - XO\|_F = \operatorname*{argmax}_{O\in\mathcal{O}(r)}\langle XO, U\rangle = \operatorname*{argmax}_{O\in\mathcal{O}(r)}\langle O, WDV^\top\rangle = WV^\top.$$

Therefore, $U^\top \bar{X} = U^\top XWV^\top = VDV^\top$ is a symmetric matrix, which implies that $H^\top \bar{X} = U^\top \bar{X} - \bar{X}^\top \bar{X}$ is also symmetric matrix.

Similarly, from formula (7.28), it is easy to obtain

$$\mathbb{E}\left[(a_i^\top H\bar{X}^\top a_i)^2\right] \leq \sigma_1\|H\|_F^2 + \mathrm{tr}^2(H^\top \bar{X}) + \|H^\top \bar{X}\|_F^2.$$

For 3) of the Claim 7.2, using the notation $\hat{H}, \tilde{H}, \hat{h}_s, \tilde{h}_s$ above, we have

$$\begin{aligned}
&\mathbb{E}\left[(a_i^\top H\bar{X}^\top a_i)^2\bar{\rho}_{i,\alpha}\right] = \mathbb{E}\left[(a_i^\top \hat{H}X^\top a_i)^2\bar{\rho}_{i,\alpha}\right]\\
&= \mathbb{E}\left[\sum_{s=1}^r \|x_s\|^2(\tilde{h}_s^\top a_i)^2 a_{i,s}^2 \cdot \prod_{t=1}^r \exp\left(-\frac{\|x_t\|^2 a_{i,t}^2}{0.99\alpha}\right)\right]\\
&\qquad + \mathbb{E}\left[\sum_{s\neq k}\|x_s\|\|x_k\|(\tilde{h}_s^\top a_i)(\tilde{h}_k^\top a_i)a_{i,s}a_{i,k} \cdot \prod_{t=1}^r \exp\left(-\frac{\|x_t\|^2 a_{i,t}^2}{0.99\alpha}\right)\right]\\
&> 0.78\sum_{s=1}^r \|x_s\|^2(2\tilde{h}_{s,s}^2 + \|\tilde{h}_s\|^2) + 0.78\sum_{s\neq k}\|x_s\|\|x_k\|(\tilde{h}_{s,s}\tilde{h}_{k,k} + \tilde{h}_{s,k}\tilde{h}_{k,s})\\
&= 0.78\mathbb{E}\left[(a_i^\top H\bar{X}^\top a_i)^2\right],
\end{aligned}$$

where the last equation follows from (7.27) and the inequality comes from the following two inequalities (7.29) and (7.30):

$$\begin{aligned}
&\mathbb{E}\Big[(\tilde{h}_s^\top a_i)^2 a_{i,s}^2 \cdot \prod_{t=1}^{r} \exp(-\frac{\|x_t\|^2 a_{i,t}^2}{0.99\alpha})\Big] \\
&= \frac{1}{\gamma\omega_s}\Big(\frac{\tilde{h}_{s,1}^2}{\omega_1} + \cdots + \frac{\tilde{h}_{s,s-1}^2}{\omega_{s-1}} + \frac{3\tilde{h}_{s,s}^2}{\omega_s} + \frac{\tilde{h}_{s,s+1}^2}{\omega_{s+1}} + \cdots + \frac{\tilde{h}_{s,r}^2}{\omega_r} + \tilde{h}_{s,r+1}^2 + \cdots + \tilde{h}_{s,n}^2\Big) \\
&\geq \frac{1}{1.102^2 \cdot \gamma}(\tilde{h}_{s,1}^2 + \cdots + \tilde{h}_{s,s-1}^2 + 3\tilde{h}_{s,s}^2 + \tilde{h}_{s,s+1}^2 + \cdots + \tilde{h}_{s,n}^2) \\
&\geq \frac{1}{1.102^2 \cdot e^{1/0.99\alpha}}(2\tilde{h}_{s,s}^2 + \|\tilde{h}_s\|^2) \\
&> 0.78(2\tilde{h}_{s,s}^2 + \|\tilde{h}_s\|^2)
\end{aligned} \tag{7.29}$$

provided $\alpha \geq 20$ and the parameters ω_k, γ are defined as follows:

$$\begin{aligned}
\omega_k &:= \frac{\|x_k\|^2}{0.495\alpha} + 1 \leq 1.102, \ \forall\, 1 \leq k \leq r, \\
\gamma &:= \sqrt{\Big(\frac{\|x_1\|^2}{0.495\alpha} + 1\Big)\Big(\frac{\|x_2\|^2}{0.495\alpha} + 1\Big)\cdots\Big(\frac{\|x_r\|^2}{0.495\alpha} + 1\Big)} \leq e^{1/0.99\alpha}
\end{aligned}$$

due to the fact that $1 + x \leq e^x$ for any $x \geq 0$ and $\|X\|_F = 1$. Similarly, for any $s \neq k, 1 \leq s, k \leq r$, we have

$$\begin{aligned}
&\mathbb{E}\left[(\tilde{h}_s^\top a_i)(\tilde{h}_k^\top a_i)a_{i,s}a_{i,k} \cdot \prod_{t=1}^{r} \exp\Big(-\frac{\|x_t\|^2 a_{i,t}^2}{0.99\alpha}\Big)\right] \\
=&\frac{\tilde{h}_{s,s}\tilde{h}_{k,k} + \tilde{h}_{s,k}\tilde{h}_{k,s}}{\gamma\omega_s\omega_k} > 0.78(\tilde{h}_{s,s}\tilde{h}_{k,k} + \tilde{h}_{s,k}\tilde{h}_{k,s}).
\end{aligned} \tag{7.30}$$

Acknowledgments. The research of the second author was supported by NSFC grant (91630203, 11688101), by Youth Innovation Promotion Association CAS, Beijing Natural Science Foundation (Z180002) and by National Basic Research Program of China (973 Program 2015CB856000).

References

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

[2] E.J. Candes, Y. C. Eldar, T. Strohmer and V. Voroninski, Phase retrieval via matrix completion, *SIAM review*, **57** (2015), 225–251.

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

[4] E.J. Candes and Y. Plan, Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements, *IEEE Transactions on Information Theory*, **57** (2011), 2342–2359.

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

[6] Y. Chen and E.J. Candes, Solving random quadratic systems of equations is nearly as easy as solving linear systems, in *Advances in Neural Information Processing Systems*, 2015, pp. 739–747.

[7] Y. Chen, Y. Chi and A.J. Goldsmith, Exact and stable covariance estimation from quadratic sampling via convex programming, *IEEE Transactions on Information Theory*, **61** (2015), 4034–4059.

[8] A. Conca, D. Edidin, M. Hering and C. Vinzant, An algebraic characterization of injectivity in phase retrieval, *Applied and Computational Harmonic Analysis*, **38** (2015), 346–356.

[9] L. Demanet and P. Hand, Stable optimizationless recovery from phaseless linear measurements, *Journal of Fourier Analysis and Applications*, **20** (2014), 199–221.

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

[11] C. Fienup and J. Dainty, Phase retrieval and image reconstruction for astronomy, *Image Recovery: Theory and Application*, **231** (1987), 275.

[12] B. Gao and Z. Xu, Phaseless recovery using the gauss–newton method, *IEEE Transactions on Signal Processing*, **65** (2017), 5885–5896.

[13] R.W. Gerchberg, A practical algorithm for the determination of phase from image and diffraction plane pictures, *Optik*, **35** (1972), 237–246.

[14] R.W. Harrison, Phase problem in crystallography, *JOSA A*, **10** (1993), 1046–1055.

[15] P. Jain, P. Netrapalli and S. Sanghavi, Low-rank matrix completion using alternating minimization, In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, ACM, 2013, 665–674.

[16] R. Kueng, H. Rauhut and U. Terstiege, Low rank matrix recovery from rank one measurements, *Applied and Computational Harmonic Analysis*, **42** (2017), 88–116.

[17] Y. Li, Y. Sun and Y. Chi, Low-rank positive semidefinite matrix recovery from corrupted rank-one measurements, *IEEE Transactions on Signal Processing*, **65** (2017), 397–408.

[18] R. Meka, P. Jain, C. Caramanis and I. S. Dhillon, Rank minimization via online learning, in *Proceedings of the 25th International Conference on Machine learning*, ACM, 2008, 656–663.

[19] R.P. Millane, Phase retrieval in crystallography and optics, *JOSA A*, **7** (1990), 394–411.

[20] P. Netrapalli, P. Jain and S. Sanghavi, Phase retrieval using alternating minimization, in *Advances in Neural Information Processing Systems*, 2013, 2796–2804.

[21] H. Rauhut and U. Terstiege, Low-rank matrix recovery via rank one tight frame measurements, *Journal of Fourier Analysis and Applications*, 2016, 1–6.

[22] B. Recht, M. Fazel and P.A. Parrilo, Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization, *SIAM review*, **52** (2010), 471–501.

[23] S. Sanghavi, R. Ward and C.D. White, The local convexity of solving systems of quadratic equations, *Results in Mathematics*, **71** (2017), 569–608.

[24] Y. Shechtman, Y. C. Eldar, O. Cohen, H.N. Chapman, J. Miao and M. Segev, Phase retrieval with application to optical imaging: a contemporary overview, *IEEE signal processing magazine*, **32** (2015), 87–109.

[25] R. Vershynin, High-dimensional probability: An introduction with applications in data science, Cambridge University Press, Cambridge, 2018.

[26] Z. Xu, The minimal measurement number for low-rank matrix recovery, *Applied and Computational Harmonic Analysis*, **44** (2018), 497–508.

[27] Q. Zheng and J. Lafferty, A convergent gradient descent algorithm for rank minimization and semidefinite programming from random linear measurements, in *Advances in Neural Information Processing Systems*, 2015, 109–117.