arXiv:1912.04085v2 [math.OC] 29 Dec 2019

LINEAR CONVERGENCE OF AN ALTERNATING POLAR DECOMPOSITION METHOD FOR LOW RANK ORTHOGONAL TENSOR APPROXIMATIONS

SHENGLONG HU AND KE YE

Abstract. Low rank orthogonal tensor approximation (LROTA) is an important problem in tensor computations and their applications. A classical and widely used algorithm is the alternating polar decomposition method (APD). In this paper, an improved version iAPD of the classical APD is proposed. For the first time, all of the following four fundamental properties are established for iAPD: (i) the algorithm converges globally and the whole sequence converges to a KKT point without any assumption; (ii) it exhibits an overall sublinear convergence with an explicit rate which is sharper than the usual $O(1/k)$ for first order methods in optimization; (iii) more importantly, it converges R-linearly for a generic tensor without any assumption; (iv) for almost all LROTA problems, iAPD reduces to APD after finitely many iterations if it converges to a local minimizer.

1. Introduction

As higher order generalizations of matrices, tensors (a.k.a. hypermatrices or multi-way arrays) are ubiquitous and inevitable in mathematical modeling and scientific computing [13, 25, 36, 47, 52]. Among numerous tensor problems studied in recent years, tensor approximation and its related topics have been becoming the main focus, see [12, 34, 36] and references therein. Applications of tensor approximation are diverse and broad, including signal processing [13], computational complexity [36], pattern recognition [2], principal component analysis [11], etc. We refer interested readers to surveys [12, 23, 34, 43] and books [25, 36, 52] for more details.

Singular value decomposition (SVD) of matrices is both a theoretical foundation and a computational workhorse for linear algebra, with applications spreading throughout scientific computing and engineering [22]. SVD of a given matrix is a *rank-one orthogonal decomposition* of the matrix [22, 28], and a truncated SVD according to the non-increasing singular values is a *low rank orthogonal approximation* of that matrix by the well-known Echart-Young theorem [19]. While a higher order tensor cannot be diagonalized by orthogonal matrices in general [36], there are several generalizations of SVD from matrices to tensors, such as higher order SVD [15], *orthogonally decomposable tensor* decompositions and approximations and their variants, see [10, 21, 32, 50, 55, 61] and references therein. In this paper, we focus on the low rank orthogonal tensor approximation (abbreviated as LROTA) problem. It is a low rank tensor approximation problem with all the factor matrices being orthonormal [12, 17, 34]. This problem is of crucial importance in applications, such as blind source separation in signal processing and statistics [11–13, 47].

2010 *Mathematics Subject Classification.* 15A18; 15A69; 65F18.

Key words and phrases. Orthogonally decomposable tensors, low rank orthogonal tensor approximation, R-linear convergence, sublinear convergence, global convergence.

In the literature, several numerical methods have been proposed to solve this problem, such as Jacobi-type methods [11], for which the tensor considered usually has a symmetric structure. Interested readers are referred to [30, 39–41, 46, 59]. A more general problem is studied, where some factor matrices are orthonormal and the rest of them are unconstrained. We denote these low rank orthogonal tensor approximation problems by LROTA-s with s the number of orthonormal factor matrices. For simplicity, LROTA denotes the problem where all the factor matrices are orthonormal. For LROTA-s, a commonly adopted algorithmic framework is the *alternating minimization method (AMM)* [6], under which the alternating polar decomposition method (APD) is proposed and widely employed [10, 61]. Under a regularity condition that all matrices in certain iterative sequence are of full rank, it is proved that every converging subsequence generated by this method for LROTA converges to a stationary point of the objective function by Chen and Saad [10] in 2009. In 2012, Uschmajew established a local convergence result under some appropriate assumptions [57]. In 2015, Wang, Chu and Yu proposed an AMM with a modified polar decomposition for LROTA-1, and established the global convergence without any further assumption for a generic tensor [61]. In 2019, Guan and Chu [24] established the global convergence for LROTA-s with general s under a similar regularity condition as [10]. Very recently, Yang proposed an epsilon-alternating least square method for solve the problem LROTA-s with general s and established its global convergence without any assumption [62] [1]. On the other hand, the special case of rank-one tensor approximation was systematically studied since the work of De Lathauwer, De Moor, and Vandewalle [15]. A higher order power method, which is essentially an application of AMM, was proposed and global convergence results were established, see [49, 58, 60]. Moreover, the convergence rate was also estimated in [29, 57, 63].

Motivated by the development of the convergence analysis of the rank-one case and the general low rank case, a fundamental question is: *is there an algorithm for LROTA such that all the favorable convergence properties in the rank-one case also hold for the general low rank case*? The answer is affirmative. Given this is true, one hopes that this algorithm should be as close as possible to the widely used classical APD, so that several questions raised in the literature can be addressed [10, 24, 61]. In this paper, we provide an affirmative answer to the question. Listed below are main contributions of this paper:

(1) we propose an improved version iAPD for the alternating polar decomposition (APD) method given in [24] for solving LROTA, and show its global convergence without any assumption;
(2) we establish an overall sublinear convergence of iAPD and present an *explicit eventual convergence rate* in terms of the dimension and the order of the underlying tensor. The derived convergence rate is sub-optimal in the sense that it is sharper than the usual convergence rate $O(1/k)$ established for first order methods in the literature [5];
(3) we prove that iAPD is linear convergent for a generic tensor without any other assumption;
(4) we also show that for almost all LROTA problems, iAPD reduces to APD after finitely many iterations if it converges to a local minimizer. In particular, this relaxes the requirement for each iterative matrix being of full rank in the literature, such as [10, 24], to a simple requirement on the limit point.

[1] Yang's paper is posted during our final preparation of this paper. We can see that we both employ the proximal technique. A difference is that the proximal correction in our algorithm is adaptive, and a theoretical investigation is also given (cf. Section 5) for the execution of the proximal correction.

Other than these, we also prove that every KKT point of LROTA is nondegenerate for a generic tensor, which might be of independent interests.

The rest of this paper is organized as follows: preliminaries on multilinear algebra, differential geometry and optimization theory that will be encountered repeatedly in the sequel are included in Section 2. In particular, the LROTA problem is stated in Section 2.6; Section 3 is devoted to the analysis of the manifold structures of the set of orthogonally decomposable tensors. In this section, the connection between KKT points of LROTA and critical points of the projection function on manifolds is established. It is shown that every KKT point of LROTA for a generic tensor is nondegenerate; the new algorithm iAPD is proposed in Section 4 and detailed convergence analysis for this algorithm is given. The overall sub-optimal sublinear convergence and generic linear convergence are also proved; Section 5 proves that for almost all LROTA problems, iAPD reduces to APD after finitely many iterations if it converges to a local minimizer; some final remarks are given in Section 6; to avoid distracting readers too much by technical details, lemmas are stated when they are needed and proofs are provided in Appendix A and B.

2. Preliminaries

2.1. **Tensors.** In this subsection, we provide a review of basic notions of tensors. Given positive integers $k \geqslant 2$ and $n_1, \dots, n_k$, the tensor space consisting of real tensors of dimension $n_1 \times \cdots \times n_k$ is denoted by $\mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_k}$. In this vector space, an inner product and hence a norm can be defined. The Hilbert-Schmidt inner product of two given tensors $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_k}$ is defined by

$$\langle \mathcal{A}, \mathcal{B} \rangle := \sum_{i_1=1}^{n_1} \cdots \sum_{i_k=1}^{n_k} a_{i_1 \dots i_k} b_{i_1 \dots i_k}.$$

The Hilbert-Schmidt norm $\|\mathcal{A}\|$ is then given by (cf. [43])

$$\|\mathcal{A}\| := \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}.$$

In particular, if $k = 2$, then an element in $\mathbb{R}^{n_1} \otimes \mathbb{R}^{n_2}$ is simply an $n_1 \times n_2$ matrix A, whose Hilbert-Schmidt norm reduces to the Frobenius norm $\|A\|_F$.

Given a positive integer r and $\lambda_1, \dots, \lambda_r \in \mathbb{R}$, we denote by $\operatorname{diag}(\lambda_1, \dots, \lambda_r)$ the diagonal tensor in $\mathbb{R}^r \otimes \cdots \otimes \mathbb{R}^r$ with the order being understood from the context. To be more precise, we have

$$(\operatorname{diag}(\lambda_1, \dots, \lambda_r))_{i_1 \dots i_k} = \begin{cases} \lambda_j, & \text{if } i_1 = \cdots = i_k = j \in \{1, \dots, r\}, \\ 0, & \text{otherwise.} \end{cases}$$

For a given positive integer k, we may regard the tensor $\operatorname{diag}(\lambda_1, \dots, \lambda_r)$ as the image of $(\lambda_1, \dots, \lambda_r)$ under the map $\operatorname{diag} : \mathbb{R}^r \to \otimes^k \mathbb{R}^r$ defined in an obvious way. We also define the map $\operatorname{Diag} : \otimes^k \mathbb{R}^r \to \mathbb{R}^r$ by taking the diagonal of a kth order tensor. By definition, $\operatorname{Diag} \circ \operatorname{diag} : \mathbb{R}^r \to \mathbb{R}^r$ is the identity map.

We define a map $\tau : \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_k} \to \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_k}$ by

$$\tau(\mathbf{x}) := \mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_k. \tag{1}$$

where $\mathbf{x}$ is a *block vector*

$$\mathbf{x} := (\mathbf{x}_1, \dots, \mathbf{x}_k) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_k} \simeq \mathbb{R}^{n_1 + \cdots + n_k} \text{ with } \mathbf{x}_i \in \mathbb{R}^{n_i} \text{ for all } i = 1, \dots, k.$$

For each $i \in \{1,\dots,k\}$, we define a map $\tau_i : \mathbb{R}^{n_1}\times\cdots\times\mathbb{R}^{n_k} \to \mathbb{R}^{n_1}\otimes\cdots\otimes\mathbb{R}^{n_{i-1}}\otimes\mathbb{R}^{n_{i+1}}\otimes\cdots\otimes\mathbb{R}^{n_k}$ by

$$\tau_i(\mathbf{x}) := \mathbf{x}_1\otimes\cdots\otimes\mathbf{x}_{i-1}\otimes\mathbf{x}_{i+1}\otimes\cdots\otimes\mathbf{x}_k, \quad \mathbf{x}\in\mathbb{R}^{n_1}\times\cdots\times\mathbb{R}^{n_k}.$$

Given a tensor $\mathcal{A}\in\mathbb{R}^{n_1}\otimes\cdots\otimes\mathbb{R}^{n_k}$ and a block vector $\mathbf{x}$ as above, $\mathcal{A}\tau(\mathbf{x})$ is defined by

$$\mathcal{A}\tau(\mathbf{x}) := \langle\mathcal{A},\tau(\mathbf{x})\rangle$$

and $\mathcal{A}\tau_i(\mathbf{x})\in\mathbb{R}^{n_i}$ is a vector defined implicitly by the relation:

$$\langle\mathcal{A}\tau_i(\mathbf{x}),\mathbf{x}_i\rangle = \mathcal{A}\tau(\mathbf{x})$$

for any block vector $\mathbf{x}$. Moreover, given k matrices $B^{(i)}\in\mathbb{R}^{m_i\times n_i}$ for $i\in\{1,\dots,k\}$, the *matrix-tensor product* $(B^{(1)},\dots,B^{(k)})\cdot\mathcal{A}$ is a tensor in $\mathbb{R}^{m_1}\otimes\cdots\otimes\mathbb{R}^{m_k}$, defined entry-wisely as

$$\big[(B^{(1)},\dots,B^{(k)})\cdot\mathcal{A}\big]_{i_1\dots i_k} := \sum_{j_1=1}^{n_1}\cdots\sum_{j_k=1}^{n_k} b^{(1)}_{i_1j_1}\dots b^{(k)}_{i_kj_k}a_{j_1\dots j_k} \tag{2}$$

for all $i_t\in\{1,\dots,m_t\}$ and $t\in\{1,\dots,k\}$.

2.2. **Stiefel manifolds.** Let $m\leqslant n$ be two positive integers and let $V(m,n)\subseteq\mathbb{R}^{n\times m}$ be the set of all $n\times m$ orthonormal matrices, i.e.,

$$V(m,n) := \{U\in\mathbb{R}^{n\times m} : U^{\mathsf{T}}U = I\},$$

where I is the $m\times m$ identity matrix. Indeed, $V(m,n)$ admits a smooth manifold structure and is called *the Stiefel manifold of orthonormal m-frames in* $\mathbb{R}^n$. In particular, if $m=n$ then $V(n,n)$ simply reduces to the orthogonal group $\mathrm{O}(n)$.

For any $A\in V(m,n)$, the *Fréchet normal cone* of $V(m,n)$ at A is defined as (cf. [53])

$$\hat{N}_{V(m,n)}(A) := \{B\in\mathbb{R}^{n\times m} \mid \langle B, C-A\rangle \leqslant o(\|C-A\|_F) \text{ for all } C\in V(m,n)\}.$$

Usually, we set $\hat{N}_{V(m,n)}(A)=\emptyset$ whenever $A\notin V(m,n)$. The *(limiting) normal cone* $N_{V(n,m)}(A)$ of $V(n,m)$ at $A\in V(n,m)$ is defined by (cf. [53])

$$N_{V(m,n)}(A) := \left\{B\in\mathbb{R}^{n\times m} : \begin{array}{c} A_k\in V(m,n), \lim\limits_{k\to\infty} A_k = A, \\ B_k\in\hat{N}_{V(m,n)}(A_k), \lim\limits_{k\to\infty} B_k = B\end{array}\right\}.$$

It is easily seen from the definition that the normal cone $N_{V(m,n)}(A)$ is always closed. The indicator function $\delta_{V(m,n)}$ of $V(m,n)$ is defined by

$$\delta_{V(m,n)}(X) := \begin{cases} 0 & \text{if } X\in V(m,n), \\ +\infty & \text{otherwise.}\end{cases}$$

Given a function $f:\mathbb{R}^n\to\mathbb{R}\cup\{\infty\}$, the *regular subdifferential* of f at $\mathbf{x}\in\mathbb{R}^n$ is defined as

$$\hat{\partial}f(\mathbf{x}) := \left\{\mathbf{v}\in\mathbb{R}^n : \liminf_{\mathbf{x}\neq\mathbf{y}\to\mathbf{x}} \frac{f(\mathbf{y})-f(\mathbf{x})-\langle\mathbf{v},\mathbf{y}-\mathbf{x}\rangle}{\|\mathbf{y}-\mathbf{x}\|} \geqslant 0\right\}$$

and the *(limiting) subdifferential* of f at $\mathbf{x}$ is defined as

$$\partial f(\mathbf{x}) := \left\{\mathbf{v}\in\mathbb{R}^n : \exists\{\mathbf{x}^k\}\to\mathbf{x} \text{ and } \{\mathbf{v}^k\}\to\mathbf{v} \text{ satisfying } \mathbf{v}^k\in\hat{\partial}f(\mathbf{x}^k) \text{ for all } k\right\}.$$

If $\mathbf{0} \in \partial f(\mathbf{x})$, then $\mathbf{x}$ is a *critical point* of f. An important fact about the normal cone $N_{V(m,n)}(A)$ and the subdifferential of the indicator function $\delta_{V(m,n)}$ of $V(m,n)$ at A is (cf. [53])

$$\partial\delta_{V(m,n)} = N_{V(m,n)}. \tag{3}$$

Note that $V(m,n)$ is a smooth manifold of dimension $mn - \frac{m(m+1)}{2}$. It follows from [53, Chapter 6.C] and [1,20] that

$$N_{V(m,n)}(A) = \hat{N}_{V(m,n)}(A) = \{AS \mid S \in \mathrm{S}^{m\times m}\},$$

where $\mathrm{S}^{m\times m} \subseteq \mathbb{R}^{m\times m}$ is the subspace of $m \times m$ symmetric matrices.

Given a matrix $B \in \mathbb{R}^{n\times m}$, the projection of B onto the normal cone of $V(m,n)$ at A is

$$\pi_{N_{V(m,n)}(A)}(B) = A(\frac{A^\intercal B + B^\intercal A}{2}).$$

The tangent space $T_{V(m,n)}(A)$ of $V(m,n)$ at a point $A \in V(m,n)$ is the orthogonal complement to the normal cone. Given a matrix $B \in \mathbb{R}^{n\times m}$, the projection of B onto the tangent space of $V(m,n)$ at a point $A \in V(m,n)$ is given by

$$\pi_{T_{V(m,n)}(A)}(B) = A\,\mathrm{skew}(A^\intercal B) + (I - AA^\intercal)B, \tag{4}$$

where $\mathrm{skew}(C) := \frac{C-C^\intercal}{2}$ is for a square matrix $C \in \mathbb{R}^{m\times m}$. A more explicit formula is given as

$$\pi_{T_{V(m,n)}(A)}(B) = (I - \frac{1}{2}AA^\intercal)(B - AB^\intercal A). \tag{5}$$

2.3. **Orthogonally decomposable tensor.** A tensor $\mathcal{A} \in \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_k}$ is called *orthogonally decomposable* (cf. [21,32,33,63]) if there exist orthonormal matrices

$$U^{(i)} = \begin{bmatrix}\mathbf{u}_1^{(i)} & \cdots & \mathbf{u}_r^{(i)}\end{bmatrix} \in V(r,n_i),\ i = 1,\dots,k$$

and numbers $\lambda_j \in \mathbb{R}$ for $1 \leqslant j \leqslant r \leqslant \min\{n_1,\dots,n_k\}$ such that

$$\mathcal{A} = \sum_{j=1}^{r} \lambda_j \mathbf{u}_j^{(1)} \otimes \cdots \otimes \mathbf{u}_j^{(k)}. \tag{6}$$

Here for each $i = 1,\dots,k$ and $j = 1,\dots,r$, the vector $\mathbf{u}_j^{(i)} \in \mathbb{R}^{n_i}$ is the j-th column vector of the orthonormal matrix $U^{(i)}$. Without loss of generality, we may assume that $\lambda_j \geqslant 0$ for all $j = 1,\dots,r$. Throughout this paper, we will always assume that $k \geqslant 3$.

Let r,k be positive integers and let $\mathbf{n} := (n_1,\dots,n_k)$ be a k-dimensional integer vector. We denote by $C(\mathbf{n},r) \subseteq \mathbb{R}^{n_1}\otimes\cdots\otimes\mathbb{R}^{n_k}$ the set of all *orthogonally decomposable tensors with rank at most* r, i.e.,

$$\begin{aligned} C(\mathbf{n},r) := \Big\{&\mathcal{A} \in \mathbb{R}^{n_1}\otimes\cdots\otimes\mathbb{R}^{n_k} : \mathcal{A} = (U^{(1)},\dots,U^{(k)})\cdot\mathrm{diag}(\lambda_1,\dots,\lambda_r),\\ &U^{(i)} \in V(r,n_i) \text{ for all } i \in \{1,\dots,k\},\ \lambda_j \in \mathbb{R} \text{ for all } j \in \{1,\dots,r\}\Big\}. \end{aligned} \tag{7}$$

We also let $D(\mathbf{n},r) \subseteq \mathbb{R}^{n_1}\otimes\cdots\otimes\mathbb{R}^{n_k}$ be the set of all *orthogonally decomposable tensors with rank* r, i.e.,

$$\begin{aligned} D(\mathbf{n},r) := \Big\{&\mathcal{A} \in \mathbb{R}^{n_1}\otimes\cdots\otimes\mathbb{R}^{n_k} : \mathcal{A} = (U^{(1)},\dots,U^{(k)})\cdot\mathrm{diag}(\lambda_1,\dots,\lambda_r),\\ &U^{(i)} \in V(r,n_i) \text{ for all } i \in \{1,\dots,k\},\ \lambda_j \neq 0 \text{ for all } j \in \{1,\dots,r\}\Big\}. \end{aligned} \tag{8}$$

Lemma 2.1 (Unique Decomposition)**.** *For each $\mathcal{A} \in D(\mathbf{n}, r)$, the rank-$r$ decomposition of $\mathcal{A}$ is unique. In particular, the orthogonal decomposition of $\mathcal{A}$ is unique.*

Proof. It follows from Kruskal's inequality [34–36] immediately. A direct proof can also be found in [63]. □

2.4. **Morse functions.** In the following, we introduce the notion of Morse functions and recall some of its basic properties. On a smooth manifold M, a smooth function $f : M \to \mathbb{R}$ is called a *Morse function* if each critical point of f on M is nondegenerate, i.e., the Hessian matrix of f at each critical point is non-singular. The following result is well-known, see for example [48, Theorem 6.6].

Lemma 2.2 (Projection is Generically Morse)**.** *Let M be a submanifold of $\mathbb{R}^n$. For a generic $\mathbf{a} = (a_1, \dots, a_n)^\intercal \in \mathbb{R}^n$, the Euclidean distance function*

$$f(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}\|^2$$

is a Morse function on M.

We will also need the following property (cf. [48, Corollary 2.3]) of nondegenerate critical points in the sequel.

Lemma 2.3. *Let M be a manifold and let $f : M \to \mathbb{R}$ be a smooth function. Nondegenerate critical points of f are isolated.*

To conclude this subsection, we briefly discuss how critical points behave under a local diffeomorphism. For this purpose, we recall that two smooth manifolds M_1 and M_2 are called *locally diffeomorphic* if there is a smooth map $\varphi : M_1 \to M_2$ such that for each point $\mathbf{x} \in M_1$ there exists a neighborhood $U \subseteq M_1$ of $\mathbf{x}$ and a neighborhood $V \subseteq M_2$ of $\varphi(\mathbf{x})$ such that the restriction $\varphi|_U : U \to V$ is a diffeomorphism [18]. In this case, the corresponding φ is called a *local diffeomorphism* between M_1 and M_2. It is clear from the definition that two locally diffeomorphic manifolds must have the same dimension. Moreover, we have the following result, whose proof can be found in [29, Proposition 5.2].

Proposition 2.4. *Let M_1 and M_2 be two locally diffeomorphic smooth manifolds and let $\varphi : M_1 \to M_2$ be the corresponding local diffeomorphism. Let $f : M_2 \to \mathbb{R}$ be a smooth function. Then $\mathbf{x} \in M_1$ is a (nondegenerate) critical point of $f \circ \varphi$ on M_1 if and only if $\varphi(\mathbf{x})$ is a (nondegenerate) critical point of f on M_2.*

When f is a smooth function on $\mathbb{R}^n$ and M is a submanifold of $\mathbb{R}^n$, we denote by ∇f the gradient of f as a function on $\mathbb{R}^n$, while we denote by $\operatorname{grad}(f)$ the Riemannian gradient of f as a function on M. In other words, $\operatorname{grad}(f)$ is the projection of ∇f to the tangent space of M.

2.5. **Kurdyka-Łojasiewicz property.** In this subsection, we will review some basic facts about the Kurdyka-Łojasiewicz property, which even holds for nonsmooth functions in general. Interested readers are referred to [3, 4, 8, 38].

Let p be an extended real-valued function and let $\partial p(\mathbf{x})$ be the set of sub-differentials of p at $\mathbf{x}$ (cf. [53]). We define $\operatorname{dom}(\partial p) := \{\mathbf{x} \colon \partial p(\mathbf{x}) \neq \emptyset\}$ and take $\mathbf{x}^* \in \operatorname{dom}(\partial p)$. If there exist some $\eta \in (0, +\infty]$, a neighborhood U of $\mathbf{x}^*$, and a continuous concave function $\varphi : [0, \eta) \to \mathbb{R}_+$, such that

(1) $\varphi(0) = 0$,

(2) φ is C^1 on $(0, \eta)$,
(3) for all $s \in (0, \eta)$, $\varphi'(s) > 0$, and
(4) for all $\mathbf{x} \in U \cap \{\mathbf{y} : p(\mathbf{x}^*) < p(\mathbf{y}) < p(\mathbf{x}^*) + \eta\}$, the Kurdyka-Łojasiewicz inequality holds

$$\varphi'(p(\mathbf{x}) - p(\mathbf{x}^*))\,\mathrm{dist}(\mathbf{0}, \partial p(\mathbf{x})) \geqslant 1,$$

then we say that p has the *Kurdyka-Łojasiewicz (abbreviated as KL) property* at $\mathbf{x}^*$. Here $\mathrm{dist}(\mathbf{0}, \partial p(\mathbf{x}))$ denotes the distance from $\mathbf{0}$ to the set $\partial p(\mathbf{x})$. If p is proper, lower semicontinuous, and has the KL property at each point of $\mathrm{dom}(\partial p)$, then p is said to be a *KL function*. Examples of KL functions include real subanalytic functions and semi-algebraic functions [9]. In this paper, semi-algebraic functions will be involved, we refer to [7] and references herein for more details on such functions. In particular, polynomial functions are semi-algebraic functions and hence KL functions. Another important fact is that the indicator function of a semi-algebraic set is also a semi-algebraic function [7, 9]. Also, a finite sum of semi-algebraic functions is again semi-algebraic. We assemble these facts to derive the following lemma which will be crucial to the analysis of the global convergence of our algorithm.

Lemma 2.5. *A finite sum of polynomial functions and indicator functions of semi-algebraic sets is a KL function.*

While KL-property is used for global convergence analysis, the Łojasiewicz inequality discussed in the rest of this subsection is for convergence rate analysis. The classical Łojasiewicz inequality for analytic functions is stated as follows (cf. [45]):

> **(Classical Łojasiewicz's gradient inequality)** If f is an analytic function with $f(\mathbf{0}) = 0$ and $\nabla f(\mathbf{0}) = \mathbf{0}$, then there exist positive constants μ, κ, and ϵ such that
>
> $$\|\nabla f(\mathbf{x})\| \geqslant \mu |f(\mathbf{x})|^{\kappa} \quad \text{for all} \quad \|\mathbf{x}\| \leqslant \epsilon.$$

As pointed out in [3, 8], it is often difficult to determine the corresponding exponent κ in Łojasiewicz's gradient inequality, and it is unknown for a general function. However, an estimate of the exponent κ in the gradient inequality were derived by D'Acunto and Kurdyka in [14, Theorem 4.2] when f is a polynomial function. We record this fundamental result in the next lemma, which will play a key role in our sublinear convergence rate analysis.

Lemma 2.6 (Łojasiewicz's Gradient Inequality for Polynomials)**.** *Let f be a real polynomial of degree d. Suppose that $f(\mathbf{0}) = 0$ and $\nabla f(\mathbf{0}) = \mathbf{0}$. There exist constants $c, \epsilon > 0$ such that for all $\|\mathbf{x}\| \leqslant \epsilon$, we have*

$$\|\nabla f(\mathbf{x})\| \geqslant c|f(\mathbf{x})|^{\kappa} \quad \textit{with} \quad \kappa = 1 - \frac{1}{d(3d-3)^{n-1}}.$$

Below is a manifold version of the Łojasiewicz gradient inequality [18].

Proposition 2.7 (Łojasiewicz's Gradient Inequality)**.** *Let M be a smooth manifold and let $g : M \to \mathbb{R}$ be a smooth function for which $\mathbf{z}^*$ is a nondegenerate critical point. Then there exists a neighborhood U in M of $\mathbf{z}^*$ such that for all $\mathbf{z} \in U$*

$$\|\operatorname{grad}(g)(\mathbf{z})\|^2 \geqslant \kappa |g(\mathbf{z}) - g(\mathbf{z}^*)|$$

for some $\kappa > 0$.

2.6. **Low rank orthogonal tensor approximation.** The problem considered in this paper can be described as follows: given a tensor $\mathcal{A} \in \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_k}$, find an orthogonally decomposable tensor $\mathcal{B} \in \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_k}$ of rank at most $r \leqslant \min\{n_1, \ldots, n_k\}$ such that the residual $\|\mathcal{A} - \mathcal{B}\|$ is minimized. More precisely, we will consider the following optimization problem:

$$\text{(LROTA(r))} \qquad \begin{array}{rl} \min & \|\mathcal{A} - (U^{(1)}, \ldots, U^{(k)}) \cdot \Upsilon\|^2 \\ \text{s.t.} & \Upsilon = \operatorname{diag}(\upsilon_1, \ldots, \upsilon_r),\ \upsilon_i \in \mathbb{R}, \\ & \big(U^{(i)}\big)^{\mathsf{T}} U^{(i)} = I \text{ for all } 1 \leqslant i \leqslant k. \end{array} \tag{9}$$

Proposition 2.8 (Maximization Equivalence)**.** *The optimization problem* (9) *is equivalent to*

$$\text{(mLROTA(r))} \qquad \begin{array}{rl} \max & \sum_{j=1}^r \Big(\big(\big(U^{(1)}\big)^{\mathsf{T}}, \ldots, \big(U^{(k)}\big)^{\mathsf{T}} \big) \cdot \mathcal{A} \Big)^2_{j\ldots j} \\ \text{s.t.} & \big(U^{(i)}\big)^{\mathsf{T}} U^{(i)} = I \textit{ for all } 1 \leqslant i \leqslant k \end{array} \tag{10}$$

in the following sense

1. *if* $(\mathbb{U}_*, \Upsilon_*) := \big((U_*^{(1)}, \ldots, U_*^{(k)}), \operatorname{diag}((\upsilon_*)_1, \ldots, (\upsilon_*)_r) \big)$ *is an optimizer of* (9) *with the optimal value* $\|\mathcal{A}\|^2 - \sum_{i=1}^r (\upsilon_*)_i^2$, *then* $\mathbb{U}_*$ *is an optimizer of* (10) *with the optimal value* $\sum_{i=1}^r (\upsilon_*)_i^2$;
2. *conversely, if* $\mathbb{U}_*$ *is an optimizer of* (10), *then* $(\mathbb{U}_*, \Upsilon_*)$ *is an optimizer of* (9) *where*
$$\Upsilon_* = \operatorname{diag}\Big(\operatorname{Diag}\Big(\big(\big(U_*^{(1)}\big)^{\mathsf{T}}, \ldots, \big(U_*^{(k)}\big)^{\mathsf{T}} \big) \cdot \mathcal{A} \Big) \Big).$$

Proof. By a direct calculation we may obtain

$$\begin{aligned} \|\mathcal{A} - (U^{(1)}, \ldots, U^{(k)}) \cdot \Upsilon\|^2 &= \|\mathcal{A}\|^2 + \sum_{i=1}^r \upsilon_i^2 - 2\langle \mathcal{A}, (U^{(1)}, \ldots, U^{(k)}) \cdot \Upsilon \rangle \\ &= \|\mathcal{A}\|^2 + \sum_{i=1}^r \upsilon_i^2 - 2\Big\langle \big(\big(U^{(1)}\big)^{\mathsf{T}}, \ldots, \big(U^{(k)}\big)^{\mathsf{T}} \big) \cdot \mathcal{A}, \Upsilon \Big\rangle \\ &= \|\mathcal{A}\|^2 + \sum_{i=1}^r \upsilon_i^2 - 2\sum_{i=1}^r \upsilon_i \Big[\big(\big(U^{(1)}\big)^{\mathsf{T}}, \ldots, \big(U^{(k)}\big)^{\mathsf{T}} \big) \cdot \mathcal{A} \Big]_{i\ldots i}. \end{aligned}$$

Note that υ_i in the minimization problem (9) is unconstrained for all $i \in \{1, \ldots, r\}$, and they are mutually independent. Thus, at an optimizer $(\mathbb{U}_*, \Upsilon_*) := \big((U_*^{(1)}, \ldots, U_*^{(k)}), \Upsilon_* \big)$ of (9), we must have

$$(\upsilon_*)_i = \Big[\big(\big(U_*^{(1)}\big)^{\mathsf{T}}, \ldots, \big(U_*^{(k)}\big)^{\mathsf{T}} \big) \cdot \mathcal{A} \Big]_{i\ldots i} \text{ for all } 1 \leqslant i \leqslant r \tag{11}$$

and the optimal value is

$$\|\mathcal{A}\|^2 - \sum_{i=1}^r (\upsilon_*)_i^2.$$

Therefore, problem (9) is equivalent to (10). □

3. KKT Points via Projection onto Manifolds

On the one hand, a numerical algorithm solving the optimization problem (9) (or equivalently its maximization reformulation (10)) is usually designed in the parameter space

$$V_{\mathbf{n},r} := V(r, n_1) \times \cdots \times V(r, n_k) \times \mathbb{R}^r.$$

See for example, [10, 24, 61, 62]. On the other hand, from a more geometric perspective, we can also regard problem (9) as the projection of a given tensor $\mathcal{A}$ onto $C(\mathbf{n}, r)$. A key ingredient in our study of problem (9) is the relation between these two viewpoints. Once such a connection is understood, we will be able to derive an algorithm in $V_{\mathbf{n},r}$ but analyse it in $C(\mathbf{n}, r)$. To be more precise, we will study both the problem of the projection

$$\begin{array}{rl} \min & \|\mathcal{A} - \mathcal{B}\|^2 \\ \text{s.t.} & \mathcal{B} \in D(\mathbf{n}, r), \end{array} \tag{12}$$

and its parametrization

$$\text{(LROTA-P)} \qquad \begin{array}{rl} \min & g(\mathbb{U}, \mathbf{x}) = \frac{1}{2}\|\mathcal{A} - (U^{(1)}, \dots, U^{(k)}) \cdot \operatorname{diag}(\mathbf{x})\|^2 \\ \text{s.t.} & \left(U^{(i)}\right)^{\mathsf{T}} U^{(i)} = I \text{ for all } 1 \leqslant i \leqslant k, \\ & \mathbf{x} \in \mathbb{R}^r_*, \end{array} \tag{13}$$

where $\mathbb{R}_* := \mathbb{R}\backslash\{0\}$.

We will first study properties of $C(\mathbf{n}, r)$ and $D(\mathbf{n}, r)$ and then discuss critical points of problem (12) in Section 3.1. KKT points of (10) and hence (13) will be discussed in Section 3.2. The connection between them will be studied in Section 3.3, in which a Łojasiwicz inequality for KKT points of (10) will be given. We refer to [18, 26, 48, 54] for basic facts of differential geometry, algebraic geometry and algebraic topology that will be used in the sequel.

3.1. **Geometry of orthogonally decomposable tensors.** Let

$$U_{\mathbf{n},r} := V(r, n_1) \times \cdots \times V(r, n_k) \times \mathbb{R}^r_*. \tag{14}$$

By the next proposition, $U_{\mathbf{n},r}$ parametrizes the manifold $D(\mathbf{n}, r)$.

Proposition 3.1. *For each positive integer $r \leqslant \min\{n_1, \dots, n_k\}$, the map*

$$\begin{aligned} \varphi_{\mathbf{n},r} : V(r, n_1) \times \cdots \times V(r, n_k) \times \mathbb{R}^r &\to C(\mathbf{n}, r), \\ (U^{(1)}, \dots, U^{(k)}, (\lambda_1, \dots, \lambda_r)) &\mapsto (U^{(1)}, \dots, U^{(k)}) \cdot \operatorname{diag}(\lambda_1, \dots, \lambda_r) \end{aligned}$$

is a surjective map and we have the following:

- *The permutation group $\mathfrak{S}_r$ acts on $V_{\mathbf{n},r}$ such that $\varphi_{\mathbf{n},r}$ is $\mathfrak{S}_r$-invariant.*
- *The inverse image $U_{\mathbf{n},r} = \varphi^{-1}_{\mathbf{n},r}(D(\mathbf{n}, r)) \subseteq V_{\mathbf{n},r}$ consists of points*
$$\left(U^{(1)}, \dots, U^{(k)}, (\lambda_1, \dots, \lambda_r)\right)$$
such that $\lambda_j \neq 0$ for all $1 \leqslant j \leqslant r$. In particular, $U_{\mathbf{n},r}$ is an open submanifold of $V_{\mathbf{n},r}$.
- *$U_{\mathbf{n},r}$ is a principal $\mathfrak{S}_r$-bundle on $D(\mathbf{n}, r)$, i.e., we have*
$$U_{\mathbf{n},r}/\mathfrak{S}_r \simeq D(\mathbf{n}, r).$$
- *$D(\mathbf{n}, r)$ is a smooth manifold of dimension*
$$d_{\mathbf{n},r} := r\left[\sum_{i=1}^{k} n_i - \frac{k(r+1)}{2} + 1\right].$$
- *$C(\mathbf{n}, r) = \bigsqcup_{t=0}^{r} D(\mathbf{n}, t)$ is an irreducible algebraic variety whose singularity locus is $\bigsqcup_{t=0}^{r-1} D(\mathbf{n}, t)$. In particular, $C(\mathbf{n}, r)$ has dimension $d_{\mathbf{n},r}$.*

Proof. We recall that $V(r,n)$ consists of all $n\times r$ matrices whose columns are orthonormal. Hence $V(r,n)$ admits an $\mathfrak{S}_r$ action by permuting columns. In other words, an element σ in $\mathfrak{S}_r$ can be written as an $r\times r$ permutation matrix P_σ, the action of $\mathfrak{S}_r$ on $V(r,n)$ is simply given by

$$\mathfrak{S}_r\times V(r,n)\to V(r,n),\quad (\sigma,U)\mapsto UP_\sigma,$$

and this induces an action

$$\begin{aligned}\mathfrak{S}_r\times(V(r,n_1)\times\cdots\times V(r,n_k)\times\mathbb{R}^r)&\to(V(r,n_1)\times\cdots\times V(r,n_k)\times\mathbb{R}^r)\\ (\sigma,(U^{(1)},\dots,U^{(k)},(\lambda_1,\dots,\lambda_r)))&\mapsto(U^{(1)}P_\sigma,\dots,U^{(k)}P_\sigma,(\lambda_{\sigma(1)},\dots,\lambda_{\sigma(r)})).\end{aligned}$$

Now it is straightforward to verify that $\varphi_{\mathbf{n},r}$ is $\mathfrak{S}_r$-invariant.

Since $D(\mathbf{n},r)$ consists of all orthogonally decomposable tensors with rank exactly r, its inverse image $U_{\mathbf{n},r}$ cannot contain a point of the form

$$(U^{(1)},\dots,U^{(k)},(\lambda_1,\dots,\lambda_r))$$

where $\lambda_j=0$ for some $1\leqslant j\leqslant r$ by Lemma 2.1. Moreover, we claim that any tensor of the form

$$\mathcal{T}=\varphi_{\mathbf{n},r}\left(U^{(1)},\dots,U^{(k)},(\lambda_1,\dots,\lambda_r)\right)=\left(U^{(1)},\dots,U^{(k)}\right)\cdot\operatorname{diag}(\lambda_1,\dots,\lambda_r)$$

where $\lambda_j\neq 0, 1\leqslant j\leqslant r$, must lie in $D(\mathbf{n},r)$. Indeed, by definition, we see that $\mathcal{T}$ has rank at most r. Moreover, by the orthogonality of column vectors of each $U^{(j)}, j=1,\dots,r$, the mode-1 matrix flattening $T^{(1)}\in\mathbb{R}^{n_1\times(n_2\cdots n_k)}$ of $\mathcal{T}$ has rank r, which implies that $\mathcal{T}$ has rank at least r and hence $\mathcal{T}$ has rank r [36,43]. This implies that $U_{\mathbf{n},r}$ is an open subset and hence an open submanifold of $V_{\mathbf{n},r}$.

We notice that $U_{\mathbf{n},r}$ admits an $\mathfrak{S}_r$-action by the restriction of that on $V_{\mathbf{n},r}$ and the fiber $\varphi_{\mathbf{n},r}^{-1}(\mathcal{T})\simeq\mathfrak{S}_r$ if $\mathcal{T}\in D(\mathbf{n},r)$. This implies that $U_{\mathbf{n},r}/\mathfrak{S}_r\simeq D(\mathbf{n},r)$.

Since $\mathfrak{S}_r$ is a finite group acting on $U_{\mathbf{n},r}$ freely, we conclude that $D(\mathbf{n},r)\simeq U_{\mathbf{n},r}/\mathfrak{S}_r$ is a smooth manifold whose dimension is

$$\dim D(\mathbf{n},r)=\dim U_{\mathbf{n},r}=\dim V_{\mathbf{n},r}=\sum_{j=1}^k\dim V(r,n_j)+\dim\mathbb{R}^r.$$

Observing that $\dim V(r,n)=r(n-r)+\binom{r}{2}$, we obtain the desired formula for $\dim D(\mathbf{n},r)$.

The fact that $C(\mathbf{n},r)$ is an algebraic variety follows directly from the definition. Since $C(\mathbf{n},r)$ is the image of the irreducible algebraic variety $V(\mathbf{n},r)$ under the map $\varphi_{\mathbf{n},r}$, we may conclude that $C(\mathbf{n},r)$ is irreducible. It is straightforward to verify that the rank of the differential $d\varphi_{\mathbf{n},r}$ drops at points in $\bigsqcup_{t=0}^{r-1}D(\mathbf{n},t)$ and this implies that $\bigsqcup_{t=0}^{r-1}D(\mathbf{n},t)$ is the singular locus of $C(\mathbf{n},r)$. ☐

We show in the next lemma that $U_{\mathbf{n},r}$ is locally diffeomorphic to $D(\mathbf{n},r)$.

Lemma 3.2 (Local Diffeomorphism)**.** *For any positive integers $n_1,\dots,n_k$ and $r\leqslant\min\{n_1,\dots,n_k\}$, the set $U_{\mathbf{n},r}$ is a smooth manifold and is locally diffeomorphic to the manifold $D(\mathbf{n},r)$.*

Proof. We recall from Proposition 3.1 that $U_{\mathbf{n},r}$ is a principle $\mathfrak{S}_r$-bundle on $D(\mathbf{n},r)$. In particular, since $\mathfrak{S}_r$ is a finite group, for any $\mathcal{T}\in D(\mathbf{n},r)$, the fiber $\varphi_{\mathbf{n},r}^{-1}(\mathcal{T})$ of the map

$$\varphi_{\mathbf{n},r}:U_{\mathbf{n},r}\to D(\mathbf{n},r)$$

consists of $r!$ points. Therefore, for a small enough neighbourhood $W \subseteq D(\mathbf{n}, r)$ of $\mathcal{T}$, the inverse image $\varphi_{\mathbf{n},r}^{-1}(W)$ is the disjoint union of $r!$ open subsets $W_1, \dots, W_{r!} \subseteq U_{\mathbf{n},r}$ and for each $j = 1, \dots, r!$, we have

$$(\varphi_{\mathbf{n},r})|_{W_j} : W_j \to W$$

is a diffeomorphism. □

By Lemma 3.2 and Proposition 2.4, problems on $D(\mathbf{n}, r)$ can be studied via problems on $U_{\mathbf{n},r}$. To that end, the tangent space of $U_{\mathbf{n},r}$ will be given at first. The following result can be checked directly, see [1, 20].

Proposition 3.3 (Tangent Space of $U_{\mathbf{n},r}$). *At any point $(\mathbb{U}, \mathbf{x}) \in U_{\mathbf{n},r}$, the tangent space of $U_{\mathbf{n},r}$ at $(\mathbb{U}, \mathbf{x})$ is*

$$\mathrm{T}_{(\mathbb{U},\mathbf{x})} U_{\mathbf{n},r} = \mathrm{T}_{U^{(1)}} V(r, n_1) \times \cdots \times \mathrm{T}_{U^{(k)}} V(r, n_k) \times \mathbb{R}^r, \tag{15}$$

where $\mathrm{T}_{U^{(i)}} V(r, n_i)$ is the tangent space of the Stiefel manifold $V(r, n_i)$ at $U^{(i)}$, which is

$$\mathrm{T}_{U^{(i)}} V(r, n_i) = \{Z \in \mathbb{R}^{n_i \times r} : (U^{(i)})^\intercal Z + Z^\intercal U^{(i)} = 0\}, \tag{16}$$

for all $i = 1, \dots, k$.

We can embed $U_{\mathbf{n},r}$ into $\mathbb{R}^{n_1 \times r} \times \cdots \times \mathbb{R}^{n_k \times r} \times \mathbb{R}^r$ in an obvious way and hence $U_{\mathbf{n},r}$ becomes an embedded submanifold of the latter. For a differentiable function $f : U_{\mathbf{n},r} \subseteq \mathbb{R}^{n_1 \times r} \times \cdots \times \mathbb{R}^{n_k \times r} \times \mathbb{R}^r \to \mathbb{R}$, a critical point $(\mathbb{U}, \mathbf{x})$ is a point at which the Riemannian gradient $\mathrm{grad}(f)(\mathbb{U}, \mathbf{x})$ of f at $(\mathbb{U}, \mathbf{x})$ is zero, which is equivalent to the fact that the projection of the Euclidean gradient $\nabla f(\mathbb{U}, \mathbf{x})$ onto the tangent space of $U_{\mathbf{n},r}$ at $(\mathbb{U}, \mathbf{x})$ is zero. More explicitly, we have the following characterization.

Lemma 3.4. *Let $A \in V(r, n)$ and let $f : V(r, n) \subseteq \mathbb{R}^{n \times r} \to \mathbb{R}$ be a smooth function. Then $\mathrm{grad}(f)(A) = 0$ if and only if*

$$\nabla f(A) = A(\nabla f(A))^\intercal A, \tag{17}$$

which is also equivalent to $\nabla f(A) = AP$ for some symmetric matrix $P \in \mathrm{S}^{r \times r}$. In particular, $A^\intercal \nabla f(A)$ is a symmetric matrix.

Proof. The proof of the first equivalence can be found in [44, Proposition 1]. For the second, we notice that from (17)

$$A^\intercal \nabla f(A) = A^\intercal A(\nabla f(A))^\intercal A = \nabla f(A)^\intercal A$$

and this proves that $A^\intercal \nabla f(A)$ is symmetric. Now if we set $P := A^\intercal \nabla f(A)$ then (17) can be written as

$$\nabla f(A) = AP^\intercal = AP.$$

Conversely, if $\nabla f(A) = AP$ for some symmetric matrix P, then (17) obviously holds by the symmetry of P and the fact that $A^\intercal A = I$. □

Let

$$g(\mathbb{U}, \mathbf{x}) := \frac{1}{2}\|\mathcal{A} - (U^{(1)}, \dots, U^{(k)}) \cdot \mathrm{diag}(\mathbf{x})\|^2 \tag{18}$$

be the objective function of (13). Since the feasible set $D(\mathbf{n}, r)$ (resp. $U_{\mathbf{n},r}$) of (12) (resp. (13)) is a smooth manifold, we may apply Proposition 2.4, Lemmas 2.2 and 3.2 to obtain the following

Proposition 3.5. *For a generic tensor $\mathcal{A}$, each critical point of the function g on $U_{\mathbf{n},r}$ is nondegenerate.*

3.2. **KKT points of LROTA.** In this subsection, we will derive the KKT system of the optimization problem (10) and study its properties.

3.2.1. *Existence.* Let $\mathbb{U} := (U^{(1)}, \dots, U^{(k)})$ be the collection of the variable matrices in (10) and for each $1 \leqslant i \leqslant k$ and $1 \leqslant j \leqslant r$, let $\mathbf{u}_j^{(i)}$ be the j-th column of the matrix $U^{(i)}$ and let

$$\mathbf{x}_j := (\mathbf{u}_j^{(1)}, \dots, \mathbf{u}_j^{(k)}) \text{ and } \mathbf{v}_j^{(i)} := \mathcal{A}\tau_i(\mathbf{x}_j).$$

For each $1 \leqslant i \leqslant k$, we define a matrix

$$V^{(i)} := \begin{bmatrix} \mathbf{v}_1^{(i)} & \dots & \mathbf{v}_r^{(i)} \end{bmatrix}, \tag{19}$$

and a diagonal matrix

$$\Lambda := \operatorname{diag}(\mathcal{A}\tau(\mathbf{x}_1), \dots, \mathcal{A}\tau(\mathbf{x}_r)). \tag{20}$$

For each $1 \leqslant j \leqslant r$, we also set

$$\lambda_j(\mathbb{U}) := \left(\left((U^{(1)})^\mathsf{T}, \dots, (U^{(k)})^\mathsf{T} \right) \cdot \mathcal{A} \right)_{j \dots j} = \mathcal{A}\tau(\mathbf{x}_j) = \langle \mathcal{A}, \mathbf{u}_j^{(1)} \otimes \cdots \otimes \mathbf{u}_j^{(k)} \rangle \tag{21}$$

Now the objective function of (10) can be re-written as

$$f(\mathbb{U}) := \sum_{j=1}^r \left(\left((U^{(1)})^\mathsf{T}, \dots, (U^{(k)})^\mathsf{T} \right) \cdot \mathcal{A} \right)_{j \dots j}^2 = \sum_{j=1}^r \lambda_j(\mathbb{U})^2. \tag{22}$$

Definition 3.6 (KKT point)**.** *Let $\mathbb{U} = (U^{(1)}, \dots, U^{(k)})$ be a feasible point of* (10)*. If there exists $\mathbb{P} = (P_1, \dots, P_k)$ where $P_i \in \mathrm{S}^{r\times r}$ for each $1 \leqslant i \leqslant k$ such that the system*

$$V^{(i)}\Lambda - U^{(i)}P_i = 0, \quad 1 \leqslant i \leqslant k. \tag{23}$$

is satisfied, then $\mathbb{U}$ is called a Karush-Kuhn-Tucker point (KKT point) and $\mathbb{P}$ is called a Lagrange multiplier associated to $\mathbb{U}$. The set of all multipliers associated to $\mathbb{U}$ is denoted by $\mathrm{M}(\mathbb{U})$*.*

It follows immediately from the system (23) that for all $1 \leqslant i \leqslant k$,

$$(U^{(i)})^\mathsf{T} V^{(i)}\Lambda = \left(V^{(i)}\Lambda\right)^\mathsf{T} U^{(i)}. \tag{24}$$

For an equality constrained optimization problem, we say that a feasible point satisfies *linear independence constraint qualification* (LICQ) if at this point all the gradients of the constraints are linearly independent.

Proposition 3.7 (LICQ)**.** *At any feasible point of the problem* (10)*, LICQ is satisfied. Thus, at any local maximizer of* (10)*, the system of KKT condition holds and* $\mathrm{M}(\mathbb{U})$ *is a singleton.*

Proof. Suppose on the contrary that LICQ is not satisfied at a feasible point $\mathbb{U} := (U^{(1)}, \dots, U^{(k)})$ of (10). Let $P_i \in \mathbb{R}^{r\times r}$ for $i = 1, \dots, k$ be the corresponding multipliers for the equality constraints in (10) such that they are not all zero. To be more precise, P_i's are defined by

$$\nabla_{\mathbb{U}} \left(\sum_{i=1}^k \left\langle \left(U^{(i)}\right)^\mathsf{T} U^{(i)} - I, P_i \right\rangle \right) = 0.$$

Aligning along the natural block partition as $\mathbb{U}$, we must have

$$\langle U^{(i)}P_i, M^{(i)} \rangle = 0, \quad M^{(i)} \in \mathbb{R}^{n_i \times r}, 1 \leqslant i \leqslant k. \tag{25}$$

Now from (25) we obtain $U^{(i)}P_i = 0$ and hence $P_i = 0$ by the orthonormality of $U^{(i)}$ for all $1 \leqslant i \leqslant k$, and this contradicts the assumption that not all P_i's are zero. Therefore, LICQ is satisfied, which implies the uniqueness of the multiplier. The rest conclusion follows from the classical theory of KKT condition [6]. □

Lemma 3.8. *A feasible point* $(\mathbb{U}, \Upsilon)$ *is a KKT point of problem* (9) *with multiplier* $\mathbb{P} := (P^{(1)}, \ldots, P^{(k)})$ *if and only if* $\mathbb{U}$ *is a KKT point of problem* (10) *with multiplier* $\mathbb{P}$ *and* $\Upsilon = \operatorname{diag}\big(\operatorname{Diag}\big(((U^{(1)})^\mathsf{T}, \ldots, (U^{(k)})^\mathsf{T}) \cdot \mathcal{A}\big)\big)$.

Proof. According to Proposition 2.8, problem (9) is equivalent to (10), from which we may obtain the desired correspondence between KKT points. □

3.2.2. *Primitive KKT points and essential KKT points.* It is possible that for some $1 \leqslant j \leqslant r$, v_j approaches to zero along iterations of an algorithm solving the problem (9). In this case, the resulting orthogonally decomposable tensor is of rank strictly smaller than r. We will discuss this degenerate case in this section.

Proposition 3.9 (KKT Reduction)**.** *Let* $\mathbb{U} = (U^{(1)}, \ldots, U^{(k)}) \in V(r, n_1) \times \cdots \times V(r, n_k)$ *be a KKT point of problem mLROTA(r) defined in* (10) *and let* $1 \leqslant j \leqslant r$ *be a fixed integer. Set*

$$\hat{\mathbb{U}} := (\hat{U}^{(1)}, \ldots, \hat{U}^{(k)}) \in V(r-1, n_1) \times \cdots \times V(r-1, n_k),$$

where for each $1 \leqslant i \leqslant k$, $\hat{U}^{(i)}$ *is the matrix obtained by deleting the* j*-th column of* $U^{(i)}$*. If* $\mathcal{A}\tau(\mathbf{x}_j) = 0$*, then* $\hat{\mathbb{U}}$ *is a KKT point of the problem mLROTA(r-1):*

$$\begin{array}{ll} \max & \| \operatorname{Diag}\big(\big(\big(U^{(1)}\big)^\mathsf{T}, \ldots, \big(U^{(k)}\big)^\mathsf{T}\big) \cdot \mathcal{A}\big) \|^2 \\ \text{s.t.} & \big(U^{(i)}\big)^\mathsf{T} U^{(i)} = I,\ U^{(i)} \in \mathbb{R}^{n_i \times (r-1)}, 1 \leqslant i \leqslant k. \end{array} \tag{26}$$

Proof. By (23), the KKT system of problem (10) is

$$V^{(i)}\Lambda = U^{(i)} P_i \text{ for all } i = 1, \ldots, k,$$

where $(P_1, \ldots, P_k) \in \mathrm{S}^{r\times r} \times \cdots \times \mathrm{S}^{r \times r}$ is the associated Lagrange multiplier. Without loss of generality, we may assume that $j = r$, which implies that the last diagonal element of Λ is zero. Thus,

$$(U^{(i)})^\mathsf{T} \begin{bmatrix} \mathbf{v}_1^{(i)} & \ldots & \mathbf{v}_{r-1}^{(i)} & \mathbf{v}_r^{(i)} \end{bmatrix} \begin{bmatrix} \hat{\Lambda} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} = P_i, \quad 1 \leqslant i \leqslant k,$$

where $\hat{\Lambda}$ is the leading $(r-1) \times (r-1)$ principal submatrix of Λ. This implies that the last column of P_i is zero. By the symmetry of P_i, we conclude that P_i is in a block diagonal form with

$$P_i = \begin{bmatrix} \hat{P}_i & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix}, \quad 1 \leqslant i \leqslant k.$$

Therefore we have

$$\begin{bmatrix} \mathbf{v}_1^{(i)} & \ldots & \mathbf{v}_{r-1}^{(i)} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\Lambda} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} = \begin{bmatrix} \hat{U}^{(i)} & \mathbf{u}_r^{(i)} \end{bmatrix} \begin{bmatrix} \hat{P}_i & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix}, \quad 1 \leqslant i \leqslant k,$$

which implies

$$\begin{bmatrix} \mathbf{v}_1^{(i)} & \ldots & \mathbf{v}_{r-1}^{(i)} \end{bmatrix} \hat{\Lambda} = \hat{U}^{(i)} \hat{P}_i, \quad 1 \leqslant i \leqslant k.$$

Consequently, we may conclude that $\hat{\mathbb{U}}$ is a KKT point of (26). □

A KKT point $\mathbb{U} = (U^{(1)}, \ldots, U^{(k)}) \in V(r, n_1) \times \cdots \times V(r, n_k)$ of problem mLROTA(r) (cf. (10)) with $\mathcal{A}\tau(\mathbf{x}_j) \neq 0$ for all $1 \leqslant j \leqslant r$ is called a *primitive KKT point* of mLROTA(r). Iteratively applying Proposition 3.9, we obtain the following

Corollary 3.10. *Let S be a proper subset of $\{1,\dots,r\}$ with cardinality $s := |S| < r$ and let $\mathbb{U} = (U^{(1)},\dots,U^{(k)}) \in V(r,n_1)\times\cdots\times V(r,n_k)$ be a KKT point of mLROTA(r). Set*

$$\hat{\mathbb{U}} := (\hat{U}^{(1)},\dots,\hat{U}^{(k)}) \in V(r-s,n_1)\times\cdots\times V(r-s,n_k),$$

where for each $1 \leqslant i \leqslant k$, $\hat{U}^{(i)}$ is obtained by deleting those columns indexed by S. If $\mathcal{A}\tau(\mathbf{x}_j) = 0$ for all $j \in S$, then $\hat{\mathbb{U}}$ is a primitive KKT point of mLROTA(r-s).

It would happen that several KKT points of mLROTA(r) reduce in this way to the same primitive KKT point of mLROTA(r-s). We call the set of such KKT points an *essential KKT point*. Therefore, there is a one to one correspondence between essential KKT points of mLROTA(r) and all primitive KKT points of mLROTA(s) for $1 \leqslant s \leqslant r$.

3.3. **Critical points are KKT points.** In this section, we will establish the relation between KKT points of problem (10) and critical points of g on the manifold $U_{\mathbf{n},r}$, which is the objective function defined in (18). To do this, we recall from (5) that the gradient of g at a point $(\mathbb{U},\mathbf{x}) \in U_{\mathbf{n},r}$ is given by

$$\operatorname{grad}_{U^{(i)}} g(\mathbb{U},\mathbf{x}) = (I - \frac{1}{2}U^{(i)}(U^{(i)})^{\mathsf{T}})\big(V^{(i)}\Gamma - U^{(i)}(V^{(i)}\Gamma)^{\mathsf{T}}U^{(i)}\big), \tag{27}$$

$$\operatorname{grad}_{\mathbf{x}} g(\mathbb{U},\mathbf{x}) = \mathbf{x} - \operatorname{Diag}\big(((U^{(1)})^{\mathsf{T}},\dots,(U^{(k)})^{\mathsf{T}})\cdot\mathcal{A}\big), \tag{28}$$

where $i = 1,\dots,k$ and $\Gamma = \operatorname{diag}(\mathbf{x})$ is the diagonal matrix formed by the vector $\mathbf{x}$.

Proposition 3.11. *A point $(\mathbb{U},\mathbf{x}) \in U_{\mathbf{n},r}$ is a critical point of g defined in* (18) *if and only if* $(\mathbb{U},\operatorname{diag}(\mathbf{x}))$ *is a KKT point of problem* (9)*.*

Proof. We recall that a critical point $(\mathbb{U},\mathbf{x}) \in U_{\mathbf{n},r}$ of g is defined by $\operatorname{grad}(g)(\mathbb{U},\mathbf{x}) = 0$. It follows from Proposition 3.3 that these critical points are defined by

$$\nabla_{U^{(i)}} g(\mathbb{U},\mathbf{x}) = U^{(i)}P_i, \quad 1 \leqslant i \leqslant r$$

where P_i is some $r\times r$ symmetric matrix and

$$\nabla_{\mathbf{x}} g(\mathbb{U},\mathbf{x}) = 0.$$

By (28), we have

$$\mathbf{x} = \operatorname{Diag}\big(((U^{(1)})^{\mathsf{T}},\dots,(U^{(k)})^{\mathsf{T}})\cdot\mathcal{A}\big),$$

and according to (27), we obtain

$$\nabla_{U^{(i)}} g(\mathbb{U},\mathbf{x}) = V^{(i)}\Lambda.$$

Therefore, by (23) a critical point of g on $U_{\mathbf{n},r}$ must come from a KKT point of problem (9). The converse is obvious and this completes the proof. $\square$

Definition 3.12 (Nondegenerate KKT Point)**.** *A KKT point* $(\mathbb{U},\Upsilon)$ *of problem* (9) *is nondegenerate if* $(\mathbb{U},\mathbf{x}) \in U_{\mathbf{n},r}$ *is a nondegenerate critical point of g with* $\operatorname{diag}(\mathbf{x}) = \Upsilon$.

Theorem 3.13 (Finite Essential Critical Points)**.** *For a generic tensor, there are only finitely many essential KKT points for problem* (10)*, and for any positive integers $r > s > 0$, a primitive KKT point of the problem mLROTA(s) corresponding to an essential KKT point of the problem mLROTA(r) is nondegenerate.*

Proof. To prove the finiteness of essential KKT points, it is sufficient to show that there are only finitely many primitive KKT points on $U_{\mathbf{n},r}$, and the finiteness follows from the layer structure of the set $V_{\mathbf{n},r}$ (or equivalently $C(\mathbf{n},r)$, cf. Proposition 3.1) and Proposition 3.9. We first recall that KKT points on $U_{\mathbf{n},r}$ are defined by (23), which is a system of polynomial equations, which implies that the set $K_{\mathbf{n},r}$ of KKT points of problem (10) on $U_{\mathbf{n},r}$ is a closed subvariety of the quasi-variety $U_{\mathbf{n},r}$. We also note that there are finitely many irreducible components of $K_{\mathbf{n},r}$ [26] and hence it suffices to prove that each irreducible component of $K_{\mathbf{n},r}$ is a singleton. Now let $Z \subseteq K_{\mathbf{n},r}$ be an irreducible component of $K_{\mathbf{n},r}$. If Z contains infinitely many points, then $\dim Z \geqslant 1$ [26]. However, each point in Z determines a critical point of the function g defined in (18) on the manifold $U_{\mathbf{n},r}$ (cf. Proposition 3.11). This implies that the set of critical points of g on $U_{\mathbf{n},r}$ has a positive dimension, which contradicts Lemma 2.3 and Proposition 3.5.

Next, by Corollary 3.10, given a non-primitive KKT point $\mathbb{U}$ of the problem mLROTA(r), we can get a primitive KKT point $\hat{\mathbb{U}}$ of problem mLROTA(s) with $s < r$ and hence we have $(\hat{\mathbb{U}}, \mathbf{x}) \in U_{\mathbf{n},s}$ where $\mathbf{x}$ is determined by $\lambda_j(\hat{\mathbb{U}})$'s. Since for a generic tensor the function g has only nondegenerate critical points on $U_{\mathbf{n},s}$ by Proposition 3.5, the second assertion follows from Proposition 3.11 and Corollary 3.10. $\square$

For simplicity, we abbreviate $\nabla_{U^{(i)}} f(\mathbb{U})$ as $\nabla_i f(\mathbb{U})$ for each $1 \leqslant i \leqslant k$. We define

$$\|\mathbb{U}\|_F^2 := \sum_{i=1}^k \|U^{(i)}\|_F^2.$$

The following result is crucial to the linear convergence analysis in the sequel.

Lemma 3.14 (Łojasiwicz's Inequality)**.** *If* $(\mathbb{U}^*, \Upsilon^*)$ *is a nondegenerate KKT point of problem* (9)*, then there exist* $\kappa > 0$ *and* $\epsilon > 0$ *such that*

$$\sum_{i=1}^k \|\nabla_i f(\mathbb{U}) - U^{(i)} \nabla_i (f(\mathbb{U}))^\intercal U^{(i)}\|_F^2 \geqslant \kappa |f(\mathbb{U}) - f(\mathbb{U}^*)| \tag{29}$$

for any $\|\mathbb{U} - \mathbb{U}^*\|_F \leqslant \epsilon$*. Here* f *is the objective function of problem* (10) *defined by* (22)*.*

Proof. Let $\delta > 0$ be the radius of the neighborhood given by Proposition 2.7. Since $(\mathbb{U}^*, \Upsilon^*)$ is a KKT point of (9), we have

$$\operatorname{Diag}(\Upsilon^*) = \operatorname{Diag}\big(((U^{(*,1)})^\intercal, \dots, (U^{(*,k)})^\intercal) \cdot \mathcal{A}\big).$$

For a given $\mathbb{U}$, let Υ be defined as

$$\operatorname{Diag}(\Upsilon) = \operatorname{Diag}\big(((U^{(1)})^\intercal, \dots, (U^{(k)})^\intercal) \cdot \mathcal{A}\big).$$

Then, there exists $\epsilon > 0$ such that

$$\|(\mathbb{U}, \Upsilon) - (\mathbb{U}^*, \Upsilon^*)\|_F \leqslant \delta \tag{30}$$

whenever $\|\mathbb{U} - \mathbb{U}^*\|_F \leqslant \epsilon$.

By Proposition 3.11, Proposition 2.7 is applicable to $(\mathbb{U}^*, \mathbf{x}^*) \in U_{\mathbf{n},r}$ for the function g. Thus, there exists $\kappa_0 > 0$ such that

$$\|\operatorname{grad}(g)(\mathbb{U}, \mathbf{x})\|^2 \geqslant \kappa_0 |g(\mathbb{U}, \mathbf{x}) - g(\mathbb{U}^*, \mathbf{x}^*)|$$

for all $\|\mathbb{U} - \mathbb{U}^*\|_F \leqslant \epsilon$, where $\mathbf{x} = \operatorname{Diag}(\Upsilon)$ is formed by the diagonal elements of Υ.

We first have

$$|g(\mathbb{U},\mathbf{x}) - g(\mathbb{U}^*,\mathbf{x}^*)| = \frac{1}{2}|f(\mathbb{U}) - f(\mathbb{U}^*)|,$$

since $f(\mathbb{U}) = \|\mathbf{x}\|^2$ and

$$\begin{aligned} g(\mathbb{U},\mathbf{x}) &= \frac{1}{2}\|\mathcal{A} - (U^{(1)},\dots,U^{(k)})\cdot \operatorname{diag}(\mathbf{x})\|^2 \\ &= \frac{1}{2}\|\mathcal{A}\|^2 - \langle \mathcal{A}, (U^{(1)},\dots,U^{(k)})\cdot \operatorname{diag}(\mathbf{x})\rangle + \frac{1}{2}\|\mathbf{x}\|^2 \\ &= \frac{1}{2}\|\mathcal{A}\|^2 - \langle \operatorname{Diag}\big(((U^{(1)})^\mathsf{T},\dots,(U^{(k)})^\mathsf{T})\cdot\mathcal{A}\big), \mathbf{x}\rangle + \frac{1}{2}\|\mathbf{x}\|^2 \\ &= \frac{1}{2}(\|\mathcal{A}\|^2 - \|\mathbf{x}\|^2). \end{aligned}$$

By (28) and the definition of $\mathbf{x}$, we also have

$$\operatorname{grad}_{\mathbf{x}} g(\mathbb{U},\mathbf{x}) = \mathbf{x} - \operatorname{Diag}\big(((U^{(1)})^\mathsf{T},\dots,(U^{(k)})^\mathsf{T})\cdot\mathcal{A}\big) = \mathbf{0}.$$

Since

$$\operatorname{grad}_{U^{(i)}} g(\mathbb{U},\mathbf{x}) = (I - \frac{1}{2}U^{(i)}(U^{(i)})^\mathsf{T})(V^{(i)}\Gamma - U^{(i)}(V^{(i)}\Gamma)^\mathsf{T}U^{(i)}),\ \text{for all } i = 1,\dots,k,$$

and

$$\nabla_i f(\mathbb{U}) = V^{(i)}\Gamma,\ \text{for all } i = 1,\dots,k,$$

where $\Gamma = \operatorname{diag}(\mathbf{x})$ is the diagonal matrix formed by the vector $\mathbf{x}$, the assertion will follow if we can show that

$$\|I - \frac{1}{2}U^{(i)}(U^{(i)})^\mathsf{T}\|_F \leqslant \kappa_1$$

is uniformly bounded by $\kappa_1 > 0$ over $\|\mathbb{U} - \mathbb{U}^*\|_F \leqslant \epsilon$. This is obviously true. The proof is then complete. □

4. iAPD Algorithm and Convergence Analysis

4.1. **Description of iAPD algorithm.** In [24], an alternating polar decomposition (APD) algorithm is proposed to solve the optimization problem (10). Algorithm 4.1 below is an improved version of the classical APD, which we call iAPD. It is an alternating polar decomposition method with adaptive proximal corrections and truncations. An iteration step in iAPD with a truncation is called a *truncation iteration*. Obviously, there are at most r truncation iterations.

Algorithm 4.1. *iAPD: Low Rank Orthogonal Tensor Approximation*

Input: a nonzero tensor $\mathcal{A} \in \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_k}$*, a positive integer* r*, and a proximal parameter* ϵ*.*

Step 0 *[Initialization]: choose* $\mathbb{U}_{[0]} := (U^{(1)}_{[0]}, \ldots, U^{(k)}_{[0]}) \in V(r, n_1) \times \cdots \times V(r, n_k)$ *such that* $f(\mathbb{U}_{[0]}) > 0$*, and a truncation parameter* $\kappa \in (0, \sqrt{f(\mathbb{U}_{[0]})/r})$*. Let* $p := 1$*.*

Step 1 *[Alternating Polar Decompositions-APD]: Let* $i := 1$*.*

Substep 1 [Polar Decomposition]: If $i > k$*, go to Step 2. Otherwise, for all* $j = 1, \ldots, r$*, let*

$$\mathbf{x}^i_{j,[p]} := (\mathbf{u}^{(1)}_{j,[p]}, \ldots, \mathbf{u}^{(i-1)}_{j,[p]}, \mathbf{u}^{(i)}_{j,[p-1]}, \mathbf{u}^{(i+1)}_{j,[p-1]}, \ldots, \mathbf{u}^{(k)}_{j,[p-1]}), \tag{31}$$

where $\mathbf{u}^{(i)}_{j,[p]}$ *is the* j*-th column of the factor matrix* $U^{(i)}_{[p]}$*.*

Compute the matrix $\Lambda^{(i)}_{[p]}$ *as*

$$\Lambda^{(i)}_{[p]} := \operatorname{diag}(\lambda^{i-1}_{1,[p]}, \ldots, \lambda^{i-1}_{r,[p]}) \text{ with } \lambda^{i-1}_{j,[p]} := \mathcal{A}\tau(\mathbf{x}^i_{j,[p]}) \text{ for } j = 1, \ldots, r, \tag{32}$$

and the matrix $V^{(i)}_{[p]}$ *as*

$$V^{(i)}_{[p]} := \begin{bmatrix} \mathbf{v}^{(i)}_{1,[p]} & \ldots & \mathbf{v}^{(i)}_{r,[p]} \end{bmatrix} \text{ with } \mathbf{v}^{(i)}_{j,[p]} := \mathcal{A}\tau_i(\mathbf{x}^i_{j,[p]}) \text{ for } j = 1, \ldots, r. \tag{33}$$

Compute the singular value decomposition of the matrix $V^{(i)}_{[p]}\Lambda^{(i)}_{[p]}$ *as*

$$V^{(i)}_{[p]}\Lambda^{(i)}_{[p]} = G^{(i)}_{[p]}\Sigma^{(i)}_{[p]}(H^{(i)}_{[p]})^\intercal,\ G^{(i)}_{[p]} \in V(r, n_i) \text{ and } H^{(i)}_{[p]} \in \mathrm{O}(r), \tag{34}$$

where the singular values $\sigma^{(i)}_{j,[p]}$*'s are ordered nonincreasingly in the diagonal matrix* $\Sigma^{(i)}_{[p]}$*. Then the polar decomposition of the matrix* $V^{(i)}_{[p]}\Lambda^{(i)}_{[p]}$ *is*

$$V^{(i)}_{[p]}\Lambda^{(i)}_{[p]} = U^{(i)}_{[p]}S^{(i)}_{[p]} \text{ with } U^{(i)}_{[p]} := G^{(i)}_{[p]}(H^{(i)}_{[p]})^\intercal,\ S^{(i)}_{[p]} := H^{(i)}_{[p]}\Sigma^{(i)}_{[p]}(H^{(i)}_{[p]})^\intercal. \tag{35}$$

Substep 2 [Proximal Correction]: If $\sigma^{(i)}_{r,[p]} < \epsilon$*, then compute the polar decomposition of the matrix* $V^{(i)}_{[p]}\Lambda^{(i)}_{[p]} + \epsilon U^{(i)}_{[p-1]}$ *as*

$$V^{(i)}_{[p]}\Lambda^{(i)}_{[p]} + \epsilon U^{(i)}_{[p-1]} = \hat{U}^{(i)}_{[p]}\hat{S}^{(i)}_{[p]} \tag{36}$$

for an orthonormal matrix $\hat{U}^{(i)}_{[p]} \in V(r, n_i)$ *and a symmetric positive semidefinite matrix* $\hat{S}^{(i)}_{[p]}$*. Update* $U^{(i)}_{[p]} := \hat{U}^{(i)}_{[p]}$*, and* $S^{(i)}_{[p]} := \hat{S}^{(i)}_{[p]}$*. Set* $i := i + 1$ *and go to Substep 1.*

Step 2 *[Truncation]: If* $\lambda^0_{j,[p+1]} = \big((U^{(i)}_{[p]})^\intercal V^{(i)}_{[p]}\big)_{jj} < \kappa$ *for some* $j \in J \subseteq \{1, \ldots, r\}$*, replace the matrices* $U^{(i)}_{[p]}$*'s by* $\hat{U}^{(i)}_{[p]}$*'s, where* $\hat{U}^{(i)}_{[p]}$ *is an* $n_i \times (r - |J|)$ *matrix formed by the columns of* $U^{(i)}_{[p]}$ *corresponding to* $\{1, \ldots, r\} \backslash J$*, for all* $i \in \{1, \ldots, k\}$*. Update* $r := r - |J|$*, and* $U^{(i)}_{[p]} := \hat{U}^{(i)}_{[p]}$ *for all* $i \in \{1, \ldots, k\}$*. Go to Step 3.*

Step 3*: Unless a termination criterion is satisfied, let* $p := p + 1$ *and go back to Step 1.*

4.2. **Properties of iAPD.** In this section, we derive some inequalities for the increments of the objective function during iterations. To do this, we define

$$\mathbb{U}_{i,[p]} := (U^{(1)}_{[p]}, \ldots, U^{(i)}_{[p]}, U^{(i+1)}_{[p-1]}, \ldots, U^{(k)}_{[p-1]}) \tag{37}$$

for each $1 \leqslant i \leqslant k$, and

$$\mathbb{U}_{[p]} := (U_{[p]}^{(1)}, \dots, U_{[p]}^{(k)}), \tag{38}$$

which is equal to $\mathbb{U}_{k,[p]} = \mathbb{U}_{0,[p+1]}$ for each $p \in \mathbb{N}$. We remark that the j-th column of a factor matrix $U_{[p]}^{(i)}$ is denoted by $\mathbf{u}_{j,[p]}^{(i)}$ for each $1 \leqslant j \leqslant r$ while the superscript i for the *block vector* $\mathbf{x}_{j,[p]}^{i}$ is not bracketed. For each $1 \leqslant j \leqslant r$ and $p \in \mathbb{N}$, we also denote

$$\lambda_{j,[p-1]}^{k} = \lambda_{j,[p]}^{0}, \tag{39}$$

where $\lambda_{j,[p]}^{i-1}$ is defined in (32) for the i-th iteration. One immediate observation is that if the p-th iteration in Algorithm 4.1 is not a truncation iteration, then the sizes of the matrices in $\mathbb{U}_{[p]}$ and those in $\mathbb{U}_{[p-1]}$ are the same. Also if the number of iterations in Algorithm 4.1 is infinite, then there is a sufficiently large N_0 such that the p-th iteration is not a truncation iteration for any $p \geqslant N_0$. The proof of the next lemma can be found in Appendix B.1.

Lemma 4.2 (Monotonicity of iAPD). *If the p-th iteration in Algorithm 4.1 is not a truncation iteration, then for each $0 \leqslant i \leqslant k-1$, we have*

$$f(\mathbb{U}_{i+1,[p]}) - f(\mathbb{U}_{i,[p]}) \geqslant \frac{\epsilon}{2}\|U_{[p]}^{(i+1)} - U_{[p-1]}^{(i+1)}\|_F^2. \tag{40}$$

Proposition 4.3 (Sufficient Decrease). *If the p-th iteration in Algorithm 4.1 is not a truncation iteration, then we have*

$$f(\mathbb{U}_{[p]}) - f(\mathbb{U}_{[p-1]}) \geqslant \frac{\epsilon}{2}\|\mathbb{U}_{[p]} - \mathbb{U}_{[p-1]}\|_F^2. \tag{41}$$

Proof. We have

$$f(\mathbb{U}_{[p]})-f(\mathbb{U}_{[p-1]}) = \sum_{i=0}^{k-1}\big(f(\mathbb{U}_{i+1,[p]})-f(\mathbb{U}_{i,[p]})\big) \geqslant \frac{\epsilon}{2}\sum_{i=0}^{k-1}\|U_{[p]}^{(i+1)}-U_{[p-1]}^{(i+1)}\|_F^2 = \frac{\epsilon}{2}\|\mathbb{U}_{[p]}-\mathbb{U}_{[p-1]}\|_F^2,$$

where the inequality follows from (40) in Lemma 4.2. □

At each truncation iteration, the number of columns of the matrices in $\mathbb{U}_{[p]}$ is decreased strictly. The first issue we have to address is that the iteration $\mathbb{U}_{[p]}$ is not vacuous, i.e., the numbers of the columns of the matrices in $\mathbb{U}_{[p]}$ are stable and positive. We have the following proposition, which is recorded for latter reference.

Proposition 4.4. *The number of columns of $U_{[p]}^{(i)}$'s will be stable at a positive integer $s \leqslant r$ and there exists N_0 such that $f(\mathbb{U}_{[p]})$ is nondecreasing for all $p \geqslant N_0$.*

Proof. Since the initial number r of columns is finite, the truncation will occur at most r times and the total decreased number of columns of matrices in $\mathbb{U}_{[p]}$ is bounded above by r. It follows from Step 2 of Algorithm 4.1 and Lemma 4.2 that if the p-th iteration is a truncation iteration and the number of columns of the matrices in $\mathbb{U}_{[p-1]}$ is decreased from r_1 to $r_2 < r_1$, then we have

$$f(\mathbb{U}_{[p]}) \geqslant f(\mathbb{U}_{[p-1]}) - (r_1 - r_2)\kappa^2.$$

By the truncation strategy in Algorithm 4.1, after all the truncation iterations, the value of the objective function will decrease at most $r\kappa^2$. Moreover, at each iteration without truncation, the value of the objective function is nondecreasing by Lemma 4.2 and $r\kappa^2 < f(\mathbb{U}_{[0]})$. Hence, $\mathbb{U}_{[p]}$ cannot be vacuous. As there can only be a finite number of truncations, there exists N_0 such that for any $p \geqslant N_0$, the p-th iteration is not a truncation iteration, and the conclusion then follows. □

Let us consider the following optimization problem

$$\max_{\mathbb{U}} h(\mathbb{U}) := f(\mathbb{U}) + \sum_{i=1}^{k} \delta_{V(r,n_i)}(U^{(i)}). \tag{42}$$

It is straightforward to verify that (42) is an unconstrained reformulation of problem (10) and h is a KL function according to Lemma 2.5. Readers can find the proof of the next lemma in Appendix B.2.

Lemma 4.5 (Subdifferential Bound). *If the $(p+1)$-st iteration is not a truncation iteration, then there exists a subdifferential $\mathbb{W}_{[p+1]} \in \partial h(\mathbb{U}_{[p+1]})$ such that*

$$\|\mathbb{W}_{[p+1]}\|_F \leqslant \sqrt{k}(2r\sqrt{r}\|\mathcal{A}\|^2 + \epsilon)\|\mathbb{U}_{[p+1]} - \mathbb{U}_{[p]}\|_F. \tag{43}$$

4.3. **Global convergence.** The following classical result can be found in [4].

Lemma 4.6 (Abstract Convergence). *Let $p : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ be a proper lower semicontinuous function and let $\{\mathbf{x}^{(k)}\} \subseteq \mathbb{R}^n$ be a sequence satisfying the following properties*

(1) there is a constant $\alpha > 0$ such that

$$p(\mathbf{x}^{(k+1)}) - p(\mathbf{x}^{(k)}) \geqslant \alpha\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|^2,$$

(2) there is a constant $\beta > 0$ and a $\mathbf{w}^{(k+1)} \in \partial p(\mathbf{x}^{(k+1)})$ such that

$$\|\mathbf{w}^{(k+1)}\| \leqslant \beta\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|,$$

(3) there is a subsequence $\{\mathbf{x}^{(k_i)}\}$ of $\{\mathbf{x}^{(k)}\}$ and $\mathbf{x}^* \in \mathbb{R}^n$ such that

$$\mathbf{x}^{(k_i)} \to \mathbf{x}^* \text{ and } p(\mathbf{x}^{(k_i)}) \to p(\mathbf{x}^*) \text{ as } i \to \infty.$$

If p has the Kurdyka-Łojasiewicz property at the point $\mathbf{x}^*$, then the whole sequence $\{\mathbf{x}^{(k)}\}$ converges to $\mathbf{x}^*$, and $\mathbf{x}^*$ is a critical point of p.

Regarding the global convergence of Algorithm 4.1, by Proposition 4.4, we can assume without loss of generality that there is no truncation iteration in the sequence $\{\mathbb{U}_{[p]}\}$ generated by Algorithm 4.1.

Proposition 4.7. *Given a sequence $\{\mathbb{U}_{[p]}\}$ generated by Algorithm 4.1, the sequence $\{f(\mathbb{U}_{[p]})\}$ increases monotonically and hence converges.*

Proof. Since the sequence $\{\mathbb{U}_{[p]}\}$ is bounded, $\{f(\mathbb{U}_{[p]})\}$ is bounded as well by the definition (cf. (22)). The convergence then follows from Proposition 4.4. □

Theorem 4.8 (Global Convergence). *Any sequence $\{\mathbb{U}_{[p]}\}$ generated by Algorithm 4.1 is bounded and converges to a KKT point of the problem (10).*

Proof. Obviously, the sequence $\{\mathbb{U}_{[p]}\}$ is bounded and the function h is continuous on the product of the Stiefel manifolds. The convergence follows from Proposition 4.3, Lemma 4.5, Lemma 4.6, and Proposition 4.7. □

4.4. **Sublinear convergence rate.** We consider the following function

$$q(\mathbb{U},\mathbb{P}) := f(\mathbb{U}) - \sum_{i=1}^{k}\langle P^{(i)}, (U^{(i)})^{\mathsf{T}}U^{(i)} - I\rangle, \tag{44}$$

which is a polynomial of degree $2k$ in $N := \sum_{i=1}^{k}(rn_i + \binom{r+1}{2})$ variables:

$$(\mathbb{U},\mathbb{P}) = (U^{(1)},\dots,U^{(k)},P^{(1)},\dots,P^{(k)}) \in \mathbb{R}^{n_1\times r}\times\cdots\times\mathbb{R}^{n_k\times r}\times \mathrm{S}^{r\times r}\times\cdots\times\mathrm{S}^{r\times r}.$$

Let

$$\tau := 1 - \frac{1}{2k(6k-3)^{N-1}}, \tag{45}$$

which is the *Łojasiewicz exponent* of the polynomial q obtained by Lemma 2.6. We suppose that $\mathbb{U}^*$ is a KKT point of (10) with the multiplier $\mathbb{P}^*$. For

$$\hat{q}(\mathbb{U},\mathbb{P}) := q(\mathbb{U},\mathbb{P}) - q(\mathbb{U}^*,\mathbb{P}^*), \tag{46}$$

we must have

$$\hat{q}(\mathbb{U}^*,\mathbb{P}^*) = 0, \quad \nabla\hat{q}(\mathbb{U}^*,\mathbb{P}^*) = 0.$$

Thus according to Lemma 2.6, there exist some $\gamma, c > 0$ such that

$$\|\nabla\hat{q}(\mathbb{U},\mathbb{P})\|_F \geqslant c|\hat{q}(\mathbb{U},\mathbb{P})|^{\tau} \text{ whenever } \|(\mathbb{U},\mathbb{P}) - (\mathbb{U}^*,\mathbb{P}^*)\|_F \leqslant \gamma.$$

Therefore,

$$\sum_{i=1}^{k}\|\nabla_i f(\mathbb{U}) - 2U^{(i)}P^{(i)}\|_F^2 \geqslant c^2(f(\mathbb{U}) - f(\mathbb{U}^*))^{2\tau} \tag{47}$$

for any feasible point $\mathbb{U}$ of (10) satisfying $\|(\mathbb{U},\mathbb{P}) - (\mathbb{U}^*,\mathbb{P}^*)\|_F \leqslant \gamma$.

Theorem 4.9 (Sublinear Convergence Rate)**.** *Let $\{\mathbb{U}_{[p]}\}$ be a sequence generated by Algorithm 4.1 for a given nonzero tensor $\mathcal{A} \in \mathbb{R}^{n_1}\otimes\cdots\otimes\mathbb{R}^{n_k}$ and let τ be defined by* (45)*. The following statements hold:*

(1) *the sequence $\{f(\mathbb{U}_{[p]})\}$ converges to f^*, with sublinear convergence rate at least $O(p^{\frac{1}{1-2\tau}})$, that is, there exist $M_1 > 0$ and $p_0 \in \mathbb{N}$ such that for all $p \geqslant p_0$*

$$f^* - f(\mathbb{U}_{[p]}) \leqslant M_1\, p^{\frac{1}{1-2\tau}}; \tag{48}$$

(2) *$\{\mathbb{U}_{[p]}\}$ converges to $\mathbb{U}^*$ globally with the sublinear convergence rate at least $O(p^{\frac{\tau-1}{2\tau-1}})$, that is, there exist $M_2 > 0$ and $p_0 \in \mathbb{N}$ such that for all $p \geqslant p_0$*

$$\|\mathbb{U}_{[p]} - \mathbb{U}^*\|_F \leqslant M_2\, p^{\frac{\tau-1}{2\tau-1}}.$$

Proof. In the following, we consider the sequence $\{\mathbb{U}_{[p]}\}$ generated by Algorithm 4.1. Let

$$P_{[p]}^{(i)} := S_{[p]}^{(i)} - \alpha_{i,[p]}I := \begin{cases} S_{[p]}^{(i)} - \epsilon I & \text{if proximal correction is executed,}\\ S_{[p]}^{(i)} & \text{otherwise,}\end{cases}$$

where $\alpha_{i,[p]} \in \{0,\epsilon\}$. We also have that

$$S_{[p]}^{(i)} = \begin{cases} (U_{[p]}^{(i)})^{\mathsf{T}}(V_{[p]}^{(i)}\Lambda_{[p]}^{(i)} + \epsilon U_{[p-1]}^{(i)}) & \text{if proximal correction is executed,}\\ (U_{[p]}^{(i)})^{\mathsf{T}}V_{[p]}^{(i)}\Lambda_{[p]}^{(i)} & \text{otherwise.}\end{cases}$$

Note that $\{\mathbb{U}_{[p]}\}$ converges by Theorem 4.8 and hence $\{V_{[p]}^{(i)}\Lambda_{[p]}^{(i)}\}$ converges. Recall that the proximal correction step is determined by singular values of the matrices $V_{[p]}^{(i)}\Lambda_{[p]}^{(i)}$'s. Thus, for sufficiently large p (say $p \geqslant p_0$), $\alpha_{i,[p]}$ will be stable for all p and $1 \leqslant i \leqslant r$. By Lemma 4.5 and (82), we have

$$\|\nabla_i f(\mathbb{U}_{[p+1]}) - 2U^{(i)}_{[p+1]} S^{(i)}_{[p+1]} + 2\alpha U^{(i)}_{[p+1]}\|_F \leqslant (r\sqrt{r}\|\mathcal{A}\|^2 + \epsilon)\|U^{(i)}_{[p+1]} - U^{(i)}_{[p]}\|_F. \tag{49}$$

Since $\{\mathbb{U}_{[p]}\}$ converges by Theorem 4.8, we see that

$$\lim_{p\to\infty} P^{(i)}_{[p]} = \lim_{p\to\infty} (U^{(i)}_{[p]})^\mathsf{T}\big(V^{(i)}_{[p]}\Lambda^{(i)}_{[p]} + \alpha U^{(i)}_{[p-1]}\big) - \alpha I = (U^{(*,i)})^\mathsf{T} V^{(*,i)}\Lambda^* = P^{(*,i)}.$$

Hence for sufficiently large p, we may conclude that

$$\|(\mathbb{U}_{[p]}, \mathbb{P}_{[p]}) - (\mathbb{U}^*, \mathbb{P}^*)\|_F \leqslant \gamma.$$

This implies

$$\begin{aligned}
c^2(f(\mathbb{U}_{[p]}) - f(\mathbb{U}^*))^{2\tau} &\leqslant \sum_{i=1}^k \|\nabla_i f(\mathbb{U}_{[p]}) - 2U^{(i)}_{[p]}P^{(i)}_{[p]}\|_F^2 \\
&\leqslant 2\sum_{i=1}^k \|\nabla_i f(\mathbb{U}_{[p+1]}) - 2U^{(i)}_{[p+1]}P^{(i)}_{[p+1]}\|_F^2 \\
&+ 2\sum_{i=1}^k \|\nabla_i f(\mathbb{U}_{[p]}) - 2U^{(i)}_{[p]}P^{(i)}_{[p]} - \big(\nabla_i f(\mathbb{U}_{[p+1]}) - 2U^{(i)}_{[p+1]}P^{(i)}_{[p+1]}\big)\|_F^2 \qquad (50) \\
&\leqslant 2(k(r\sqrt{r}\|\mathcal{A}\|^2 + \epsilon)^2 + L)\|\mathbb{U}_{[p+1]} - \mathbb{U}_{[p]}\|_F^2 \\
&\leqslant (k(r\sqrt{r}\|\mathcal{A}\|^2 + \epsilon)^2 + L)\epsilon(f(\mathbb{U}_{[p+1]}) - f(\mathbb{U}_{[p]}),
\end{aligned}$$

where the first inequality follows from (47), the third from (49) and the fact that the function in (50) is Lipschitz continuous since $\alpha_{i,[p]}$ is stable for sufficiently large p, and the last one follows from Proposition 4.3. Here L is the Lipschitz constant of the function in (50) on the product of Stiefel manifolds.

If we set $\beta_p := f(\mathbb{U}^*) - f(\mathbb{U}_{[p]})$, then we have

$$\beta_p - \beta_{p+1} \geqslant M\beta_p^{2\tau}$$

for some constant $M > 0$, from which we can show

$$\beta_{p+1}^{1-2\tau} - \beta_p^{1-2\tau} \geqslant (2\tau - 1)M.$$

Thus,

$$\beta_p^{1-2\tau} \geqslant M(2\tau - 1)(p - p_0) + \beta_{p_0}^{1-2\tau}$$

and the conclusion follows since $\tau \in (0, 1)$. For a more detailed analysis on the sequence $\{\beta_p\}$, we refer readers to [29, Section 3.4]. □

We remark that the convergence rate in (48) is faster than the classical $O(1/p)$ for first order methods in optimization [5], while the optimal rate is $O(1/p^2)$ for convex problems by the celebrated work of Nesterov [51].

4.5. **Linear convergence.** In this section, we will establish the linear convergence of Algorithm 4.1. The proof of the next lemma is available in Appendix B.3.

Lemma 4.10 (Relative Error)**.** *There exists a constant $\gamma > 0$ such that*

$$\|\nabla_i f(\mathbb{U}_{[p+1]}) - U^{(i)}_{[p+1]}(\nabla_i f(\mathbb{U}_{[p+1]}))^\mathsf{T} U^{(i)}_{[p+1]}\|_F \leqslant \gamma \|U^{(i)}_{[p]} - U^{(i)}_{[p+1]}\|_F$$

for all $1 \leqslant i \leqslant k$ and $p \in \mathbb{N}$.

Theorem 4.11 (Linear Convergence Rate)**.** *Let $\{\mathbb{U}_{[p]}\}$ be a sequence generated by Algorithm 4.1 for a given nonzero tensor $\mathcal{A} \in \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_k}$. If $\mathbb{U}_{[p]} \to \mathbb{U}^*$ with $\mathbb{U}^*$ a nondegenerate KKT point of* (10)*, then the whole sequence $\{\mathbb{U}_{[p]}\}$ converges R-linearly to $\mathbb{U}^*$.*

Proof. By Theorem 4.8, the sequence $\{\mathbb{U}_{[p]}\}$ converges globally to $\mathbb{U}^*$, which together with $\mathbf{x}^* := \operatorname{Diag}\big(((U^{(*,1)})^\mathsf{T}, \ldots, (U^{(*,k)})^\mathsf{T}) \cdot \mathcal{A}\big)$ is a nondegenerate critical point of the function g on $U_{\mathbf{n},r}$. Hence for a sufficiently large p, Lemma 3.14 implies that

$$\sum_{i=1}^{k} \|\nabla_i f(\mathbb{U}_{[p]}) - U^{(i)}_{[p]} \nabla_i (f(\mathbb{U}_{[p]}))^\mathsf{T} U^{(i)}_{[p]}\|_F^2 \geqslant \kappa |f(\mathbb{U}_{[p]}) - f(\mathbb{U}^*)|.$$

On the other hand, by Lemma 4.10, we have

$$\sum_{i=1}^{k} \|\nabla_i f(\mathbb{U}_{[p]}) - U^{(i)}_{[p]} \nabla_i (f(\mathbb{U}_{[p]}))^\mathsf{T} U^{(i)}_{[p]}\|_F^2 \leqslant k\gamma^2 \|\mathbb{U}_{[p]} - \mathbb{U}_{[p-1]}\|_F^2.$$

Thus,

$$\begin{aligned} f(\mathbb{U}_{[p]}) - f(\mathbb{U}_{[p-1]}) &\geqslant \frac{\epsilon}{2} \|\mathbb{U}_{[p]} - \mathbb{U}_{[p-1]}\|_F^2 \\ &\geqslant \frac{\kappa\epsilon}{2k\gamma^2} (f(\mathbb{U}^*) - f(\mathbb{U}_{[p]})), \end{aligned}$$

where the first inequality follows from Proposition 4.3, and the second follows from the preceding two inequalities and Proposition 4.7. Therefore, for a sufficiently large p, we have

$$f(\mathbb{U}^*) - f(\mathbb{U}_{[p]}) \leqslant \frac{2k\gamma^2}{2k\gamma^2 + \kappa\epsilon} \big(f(\mathbb{U}^*) - f(\mathbb{U}_{[p-1]})\big), \tag{51}$$

which establishes the local Q-linear convergence of the sequence $\{f(\mathbb{U}_{[p]})\}$. Consequently, we have

$$\begin{aligned} \|\mathbb{U}_{[p]} - \mathbb{U}_{[p-1]}\|_F &\leqslant \sqrt{\frac{2}{\epsilon}} \sqrt{f(\mathbb{U}_{[p]}) - f(\mathbb{U}_{[p-1]})} \\ &\leqslant \sqrt{\frac{2}{\epsilon}} \sqrt{f(\mathbb{U}^*) - f(\mathbb{U}_{[p-1]})} \\ &\leqslant \sqrt{\frac{2}{\epsilon}} \left[\sqrt{\frac{2k\gamma^2}{2k\gamma^2 + \kappa\epsilon}}\right]^{p-1} \sqrt{f(\mathbb{U}^*) - f(\mathbb{U}_{[0]})}, \end{aligned}$$

which implies that

$$\sum_{p=p_0}^{\infty} \|\mathbb{U}_{[p]} - \mathbb{U}_{[p-1]}\|_F < \infty$$

for any sufficiently large positive integer p_0. As $\mathbb{U}_{[p]} \to \mathbb{U}^*$, we have

$$\|\mathbb{U}_{[p]} - \mathbb{U}^*\|_F \leqslant \sum_{s=p}^{\infty} \|\mathbb{U}_{[s+1]} - \mathbb{U}_{[s]}\|_F.$$

Hence, we obtain

$$\|\mathbb{U}_{[p]} - \mathbb{U}^*\|_F \leqslant \sqrt{\frac{2}{\epsilon}}\sqrt{f(\mathbb{U}^*) - f(\mathbb{U}_{[0]})}\frac{1}{1-\sqrt{\frac{2k\gamma^2}{2k\gamma^2+\kappa\epsilon}}}\Big[\sqrt{\frac{2k\gamma^2}{2k\gamma^2+\kappa\epsilon}}\Big]^p,$$

which is the claimed R-linear convergence of the sequence $\{\mathbb{U}_{[p]}\}$ and this completes the proof. □

The following result follows from Theorems 4.11 and 3.13.

Theorem 4.12 (Generic Linear Convergence). *If $\{\mathbb{U}_{[p]}\}$ is a sequence generated by Algorithm 4.1 for a generic tensor $\mathcal{A} \in \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_k}$, then the sequence $\{\mathbb{U}_{[p]}\}$ converges R-linearly to a KKT point of* (10)*.*

5. Discussions on Proximal Correction and Truncation

In this section, we will carry out a further study of proximal corrections and truncation iterations in Algorithm 4.1. We will prove that if we make an appropriate assumption on the limiting point, then the truncation iteration is unnecessary and proximal corrections are only needed for finitely many times. Thus, our iAPD reduces to the classical APD proposed in [24] after finitely many iterations. Remarkably, the assumption on the whole iteration sequence (cf. [24, Assumption A]) is vastly relaxed to a requirement on the limiting point. Together with conclusions in Section 3 about KKT points, our results in this section can shed some light on the further understanding of APD and iAPD.

5.1. **Proximal correction.**

In this subsection, we will prove that in most situations, the proximal correction in Algorithm 4.1 is unnecessary. Before we proceed, we introduce the notion of *regular KKT point*.

Definition 5.1 (Regular KKT Point). *A KKT point $\mathbb{U} := (U^{(1)}, \dots, U^{(k)}) \in \mathbb{R}^{n_1\times r} \times \cdots \times \mathbb{R}^{n_k\times r}$ of* (10) *is called a regular KKT point if the matrix $V^{(i)}\Lambda$ (cf.* (19) *and* (20)*) is of rank at least* $\min\{r, n_i - 1\}$ *for each* $1 \leqslant i \leqslant k$*.*

The requirement for a regular KKT point in Definition 5.1 is natural. The matrix $U^{(i)}$ is orthonormal and hence it is of full rank, for each $1 \leqslant i \leqslant k$. On the other hand, these matrices are polar orthonormal factor matrices of $V^{(i)}\Lambda$'s by Algorithm 4.1. If for some $1 \leqslant i \leqslant k$, the matrix $V^{(i)}\Lambda$ is of defective rank, then the best rank r approximation of the i-th factor matrix is not unique. The case that $r = n_i$ is an exceptional case, see Lemma A.8. With Lemma A.8, we have a revised proximal correction step, which is described in Algorithm 5.2.

Algorithm 5.2. *Revisited Proximal Step*

$\tau > \epsilon$ is a given constant.

Substep 2 [Revised Proximal Correction]: If $\sigma^{(i)}_{r,[p]} < \epsilon$, then consider the following two cases.

Case (i) *If $r = n_i$ and $\sigma^{(i)}_{r-1,[p]} \geqslant \tau$, then define a vector*

$$\hat{\mathbf{g}}^{(i)}_r := \begin{cases} -\mathbf{g}^{(i)}_r & \text{if } \langle \mathbf{g}^{(i)}_r, (U^{(i)}_{[p-1]} H^{(i)}_{[p]})_r \rangle < 0 \\ \mathbf{g}^{(i)}_r & \text{otherwise} \end{cases}$$

where $\mathbf{g}^{(i)}_r$ is the r-th column of the matrix $G^{(i)}_{[p]}$ and similar for $(U^{(i)}_{[p-1]} H^{(i)}_{[p]})_r$. Form a matrix $\hat{G}^{(i)}_{[p]}$ from $G^{(i)}_{[p]}$ by replacing the last column with $\hat{\mathbf{g}}^{(i)}_r$. Let

$$U^{(i)}_{[p]} := \hat{G}^{(i)}_{[p]} (H^{(i)}_{[p]})^\intercal \text{ and } S^{(i)}_{[p]} := (U^{(i)}_{[p]})^\intercal V^{(i)}_{[p]} \Lambda^{(i)}_{[p]}. \tag{52}$$

Case (ii) *For the other cases, compute the polar decomposition of the matrix $V^{(i)}_{[p]} \Lambda^{(i)}_{[p]} + \epsilon U^{(i)}_{[p-1]}$ as*

$$V^{(i)}_{[p]} \Lambda^{(i)}_{[p]} + \epsilon U^{(i)}_{[p-1]} = \hat{U}^{(i)}_{[p]} \hat{S}^{(i)}_{[p]}$$

for an orthonormal matrix $\hat{U}^{(i)}_{[p]} \in V(r, n_i)$ and a symmetric positive semidefinite matrix $\hat{S}^{(i)}_{[p]}$. Update $U^{(i)}_{[p]} := \hat{U}^{(i)}_{[p]}$ and $S^{(i)}_{[p]} := \hat{S}^{(i)}_{[p]}$.

Set $i := i + 1$ and go to Substep 1.

If we replace the Substep 2 in Algorithm 4.1 by the revised version described in Algorithm 5.2, then the sequence $\{\mathbb{U}_{[p]}\}$ still has the sufficient decreasing property. This is the content of the following lemma.

Lemma 5.3 (Revised Version)**.** *Suppose that $\tau > \epsilon > 0$. For any $p \in \mathbb{N}$ such that the p-th iteration is not a truncation iteration, we have*

$$f(\mathbb{U}_{i+1,[p]}) - f(\mathbb{U}_{i,[p]}) \geqslant \frac{1}{2} \min\{\epsilon, \tau - \epsilon\} \|U^{(i+1)}_{[p]} - U^{(i+1)}_{[p-1]}\|^2_F, \quad 0 \leqslant i \leqslant k-1. \tag{53}$$

Moreover, the matrix $S^{(i)}_{[p]}$ defined in (52) *is symmetric.*

Proof. The matrix $S^{(i)}_{[p]}$ is obviously symmetric by a direct calculation. For the decreasing property, it is sufficient to prove the result for Case (i) in Algorithm 5.2. We suppose that at iteration p and i, Case (i) of Algorithm 5.2 is executed. In this case, $\mathbf{g}^{(i)}_r$ is totally determined (up to sign) by the first $(r-1)$ columns of the matrix $G^{(i)}_{[p]}$. It follows from Lemma A.8 and the choice of $\hat{\mathbf{g}}^{(i)}_r$ that

$$\begin{aligned} \|\hat{\mathbf{g}}^{(i)}_r - (U^{(i)}_{[p-1]} H^{(i)}_{[p]})_r\| &= \min\{\|\mathbf{g}^{(i)}_r - (U^{(i)}_{[p-1]} H^{(i)}_{[p]})_r\|, \|\mathbf{g}^{(i)}_r + (U^{(i)}_{[p-1]} H^{(i)}_{[p]})_r\|\} \\ &\leqslant \|(\hat{G}^{(i)}_{[p]})_1 - (U^{(i)}_{[p-1]} H^{(i)}_{[p]})_1\|_F \\ &= \|(G^{(i)}_{[p]})_1 - (U^{(i)}_{[p-1]} H^{(i)}_{[p]})_1\|_F, \end{aligned} \tag{54}$$

where $(A)_1$ represents the $n_i \times (r-1)$ submatrix formed by the first $(r-1)$ columns of a given $n_i \times r$ matrix A. Here since $\hat{G}^{(i)}_{[p]}$ is defined by replacing the last column of $G^{(i)}_{[p]}$ by $\hat{\mathbf{g}}^{(i)}_r$, we have $(\hat{G}^{(i)}_{[p]})_1 = (G^{(i)}_{[p]})_1$, which immediately implies (54).

We then have

$$\begin{aligned}\sum_{j=1}^{r}\lambda_{j,[p]}^{i-1}(\lambda_{j,[p]}^{i}-\lambda_{j,[p]}^{i-1}) &= \operatorname{Tr}((U_{[p]}^{(i)})^{\mathsf{T}}V_{[p]}^{(i)}\Lambda_{[p]}^{(i)}) - \operatorname{Tr}((U_{[p-1]}^{(i)})^{\mathsf{T}}V_{[p]}^{(i)}\Lambda_{[p]}^{(i)})\\
&= \langle G_{[p]}^{(i)}\Sigma_{[p]}^{(i)}(H_{[p]}^{(i)})^{\mathsf{T}}, U_{[p]}^{(i)} - U_{[p-1]}^{(i)}\rangle\\
&= \langle G_{[p]}^{(i)}\Sigma_{[p]}^{(i)}, \hat{G}_{[p]}^{(i)} - U_{[p-1]}^{(i)}H_{[p]}^{(i)}\rangle\\
&\geqslant \langle (G_{[p]}^{(i)})_1\tilde{\Sigma}_{[p]}^{(i)}, (\hat{G}_{[p]}^{(i)})_1 - (U_{[p-1]}^{(i)}H_{[p]}^{(i)})_1\rangle - \epsilon|\langle \hat{\mathbf{g}}_r^{(i)}, \hat{\mathbf{g}}_r^{(i)} - (U_{[p-1]}^{(i)}H_{[p]}^{(i)})_r\rangle|\\
&\geqslant \tau\|(\hat{G}_{[p]}^{(i)})_1 - (U_{[p-1]}^{(i)}H_{[p]}^{(i)})_1\|_F^2 - \epsilon\|\hat{\mathbf{g}}_r^{(i)} - (U_{[p-1]}^{(i)}H_{[p]}^{(i)})_r\|^2\\
&\geqslant (\tau-\epsilon)\|(\hat{G}_{[p]}^{(i)})_1 - (U_{[p-1]}^{(i)}H_{[p]}^{(i)})_1\|_F^2\\
&\geqslant \frac{1}{2}(\tau-\epsilon)\|\hat{G}_{[p]}^{(i)} - U_{[p-1]}^{(i)}H_{[p]}^{(i)}\|_F^2\\
&= \frac{1}{2}(\tau-\epsilon)\|U_{[p]}^{(i)} - U_{[p-1]}^{(i)}\|_F^2, \end{aligned}\tag{55}$$

where $\tilde{\Sigma}_{[p]}^{(i)}$ is the $(r-1)\times(r-1)$ leading principal submatrix of $\Sigma_{[p]}^{(i)}$, the first inequality follows from $\sigma_{r,[p]}^{(i)} < \epsilon$, the second inequality follows from $\sigma_{r-1,[p]}^{(i)} \geqslant \tau$, the last two inequalities both follow from (54). With (55), the rest of the proof is the same as that of Lemma 4.2 and the conclusion follows. □

With Lemma 5.3, all the convergence results established in Section 4 hold as well.

Theorem 5.4 (Regular KKT point)**.** *Let $\mathcal{A}$ be a generic tensor. If $\mathbb{U}$ is a local maximizer of problem* (10) *where each entry of* $\operatorname{Diag}(\Upsilon)$ *is nonzero, then* $(\mathbb{U}, \operatorname{Diag}(\Upsilon))$ *is a nondegenerate critical point of g defined in* (18) *and $\mathbb{U}$ is a regular KKT point of problem* (10)*.*

Proof. If $\operatorname{Diag}(\Upsilon)$ is a vector with all nonzero components, then we must have $(\mathbb{U}, \operatorname{Diag}(\Upsilon)) \in U_{\mathbf{n},r}$. Moreover, Proposition 3.11 implies that $(\mathbb{U}, \operatorname{Diag}(\Upsilon))$ is a critical point of g and since g is a Morse function on $U_{\mathbf{n},r}$ for a generic tensor $\mathcal{A}$ by Proposition 3.5, it is actually a nondegenerate critical point. According to Lemma 2.3, we may conclude that $(\mathbb{U}, \operatorname{Diag}(\Upsilon))$ is isolated, i.e., g has no other non-degenerate critical point near $(\mathbb{U}, \operatorname{Diag}(\Upsilon))$.

In the following, we will prove that the matrix $V^{(i)}\Lambda$ defined by (19) and (20) is of rank at least $\min\{r, n_i - 1\}$ for all $i = 1, \ldots, k$. Suppose on the contrary that there exists some $i \in \{1, \ldots, k\}$ such that the matrix $V^{(i)}\Lambda$ has rank $s < \min\{r, n_i - 1\}$. We consider the singular value decomposition of $V^{(i)}\Lambda$

$$V^{(i)}\Lambda = U\Sigma V^{\mathsf{T}}$$

with $U := [U_1\ U_2] \in V(r, n_i)$, $U_1 \in \mathbb{R}^{n_i\times s}$, $\Sigma \in \mathbb{R}^{r\times r}$, $V = [V_1\ V_2] \in \mathrm{O}(r)$ and $V_1 \in \mathbb{R}^{r\times s}$. Hence

$$V^{(i)}\Lambda = (UV^{\mathsf{T}})(V\Sigma V^{\mathsf{T}})$$

is a polar decomposition of $V^{(i)}\Lambda$ with the polar orthonormal factor matrix UV^{T}. Since the rank of $V^{(i)}\Lambda$ is $s < \min\{r, n_i - 1\}$, the polar decomposition of the matrix $V^{(i)}\Lambda$ is not unique and has the form

$$V^{(i)}\Lambda = P(V\Sigma V^{\mathsf{T}}),$$

where

$$P = U_1V_1^{\mathsf{T}} + U_2QV_2^{\mathsf{T}} \in V(r, n_i) \text{ for some } Q \in \mathrm{O}(r-s).$$

Since $\min\{r, n_i - 1\} > s$, we must have $n_i - s \geqslant 2$ and this implies that U_2 can be chosen from the following set

$$C := \{W_2 \in \mathbb{R}^{n_i \times (r-s)} \colon [U_1\ W_2] \in V(r, n_i)\}.$$

Since U_1 is a fixed element in $V(s, n_i)$, C is isomorphic to $V(r-s, n_i - s)$ and hence $C \subseteq \mathbb{R}^{n_i \times (r-s)}$ is an irreducible closed subvariety of dimension

$$\dim C = \frac{1}{2}(r-s)\left((n_i - r) + (n_i - s - 1)\right) \geqslant 1.$$

Therefore, in any small neighborhood of P, there exists an orthonormal matrix $\tilde{P}$ such that $\tilde{P}$ also gives a polar orthonormal factor matrix of $V^{(i)}\Lambda$. Now if we fix the other $U^{(i)}$'s and Υ, then

$$\begin{aligned} g((U^{(1)}, \dots, U^{(i-1)}, \tilde{P}, U^{(i+1)}, \dots, U^{(k)}), \operatorname{Diag}(\Lambda)) &= \|\mathcal{A}\|^2 - 2\langle V^{(i)}\Lambda, \tilde{P}\rangle + \|\Upsilon\|^2 \\ &= \|\mathcal{A}\|^2 - 2\langle V^{(i)}\Lambda, P\rangle + \|\Upsilon\|^2, \end{aligned}$$

where the second equality follows from the fact that $\tilde{P}$ is also a polar orthonormal factor matrix of $V^{(i)}\Lambda$. Therefore, g is constant at such a point

$$((U^{(1)}, \dots, U^{(i-1)}, \tilde{P}, U^{(i+1)}, \dots, U^{(k)}), \operatorname{Diag}(\Upsilon)).$$

Since $(\mathbb{U}, \Upsilon)$ is a local minimizer of (9) (or equivalently, $\mathbb{U}$ is a local maximizer of (10)), we may conclude that such a point is also a local minimizer of (9). In particular, each such point corresponds to a critical point of g on $U_{\mathbf{n},r}$ by Proposition 3.7 and Proposition 3.11. However, this contradicts to the fact that $(\mathbb{U}, \operatorname{Diag}(\Upsilon))$ is an isolated critical point of g on $U_{\mathbf{n},r}$. □

Corollary 5.5. *For a generic tensor $\mathcal{A}$, there exist $\tau > \epsilon > 0$ such that if Substep 2 in Algorithm 4.1 is replaced by Algorithm 5.2, then the proximal step (i.e., Case (ii) in Algorithm 5.2) will only be executed finitely many times if the algorithm converges to a local maximizer of* (10).

Proof. For a generic tensor, there are finitely many essential KKT points whose corresponding primitive KKT points are all nondegenerate by Theorem 3.13. Therefore, there exists a constant $\tau > 0$ such that it is strictly smaller than the smallest positive singular values of all $V^{(i)}\Lambda$'s determined by these primitive KKT points.

We take $0 < \epsilon < \tau$ and let $\{\mathbb{U}_{[p]}\}$ be a sequence generated by the modified algorithm which converges to $\mathbb{U}^*$. Note that the convergence is guaranteed by Lemma 5.3 and results in Section 4. Let Λ^* be the limit of $\Lambda^{(i)}_{[p]}$ defined in (32) and $V^{(*,i)}$ the limit of $V^{(i)}_{[p]}$ defined in (33). The truncation iteration ensures that $\operatorname{Diag}(\Lambda^*)$ is a vector with each component nonzero. Thus, by Theorem 5.4, the limit point $\mathbb{U}^*$ is a regular KKT point of (10). Consequently, the rank of the matrix $V^{(*,i)}\Lambda^*$ is either of full rank and hence the proximal correction step will not be executed by the choice of ϵ and τ, or the rank of the matrix $V^{(*,i)}\Lambda^*$ is of rank $(r-1)$ when $r = n_i$, in which case Case (i) in Algorithm 5.2 will be executed by the choice of ϵ and τ. Therefore, after finitely many iterations, Case (ii) in Algorithm 5.2 will not be executed. □

5.2. **Truncation.** In this subsection, we will prove that for almost all LROTA problems, local minimizers of (9) are actually contained in the manifold $D(\mathbf{n}, r)$. Therefore, if Algorithm 4.1 converges to a local minimizer of (9), we can choose a suitable $\kappa > 0$ such that the truncation step (i.e., Step 2) in Algorithm 4.1 is unnecessary.

Theorem 5.6. *If the sequence* $\mathbf{n} = (n_1, \ldots, n_k)$ *and the positive integer* $r \leqslant \min\{n_1, \ldots, n_k\}$ *satisfies the relation*

$$d_{\mathbf{n},r-1} < \prod_{i=1}^{k}(n_i - r + 1), \tag{56}$$

where $d_{\mathbf{n},r-1} := (r-1)\left[\sum_{i=1}^{k} n_i - \frac{kr}{2} + 1\right]$, *then for a generic* $\mathcal{A} \in \mathbb{R}^{n_1\times\cdots\times n_k}$, *each local minimizer of problem* (9) *is of the form* $(U^{(1)}, \ldots, U^{(k)}, (\lambda_1, \ldots, \lambda_r)) \in \mathrm{V}(r, n_1)\times\cdots\times\mathrm{V}(r, n_k)\times\mathbb{R}^r$ *such that*

$$(U^{(1)}, \ldots, U^{(k)}) \cdot \operatorname{diag}(\lambda_1, \ldots, \lambda_r) \in D(\mathbf{n}, r).$$

Proof. We consider the subset $Z \subseteq \mathbb{R}^{n_1\times\cdots\times n_k}$ consisting of tensors of the form

$$(V^{(1)}, \ldots, V^{(k)}) \cdot \operatorname{diag}(\mu_1, \ldots, \mu_{r-1}) + \mathcal{X}.$$

Here $V^{(i)} \in \mathrm{V}(r-1, n_i)$, $\mu_j \in \mathbb{R}$, and $\mathcal{X}$ is a linear combination of decomposable tensors $\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_k$ where for each $i = 1, \ldots, k$,

(1) $\mathbf{v}_i \in \mathbb{R}^{n_i}$ is a unit norm vector;
(2) if $\mathbf{v}_i$ is not a column vector of $V^{(i)}$, then $\mathbf{v}_i^{\intercal} V^{(i)} = \mathbf{0}$;
(3) and there exists some $1 \leqslant j \leqslant k$ such that $\mathbf{v}_j$ is a column vector of $V^{(j)}$.

We notice that

$$(V^{(1)}, \ldots, V^{(k)}) \cdot \operatorname{diag}(\mu_1, \ldots, \mu_{r-1}) \in C(\mathbf{n}, r-1)$$

and $\mathcal{X}$ is contained in a vector space of dimension at most

$$\prod_{i=1}^{k} n_i - \prod_{i=1}^{k}(n_i - r + 1).$$

This implies that the dimension of Z is bounded above by

$$\dim C(\mathbf{n}, r-1) + \left(\prod_{i=1}^{k} n_i - \prod_{i=1}^{k}(n_i - r + 1)\right) = d_{\mathbf{n},r-1} + \left(\prod_{i=1}^{k} n_i - \prod_{i=1}^{k}(n_i - r + 1)\right),$$

since $\dim C(\mathbf{n}, r-1) = d_{\mathbf{n},r-1}$ by Proposition 3.1. In particular, we have

$$\dim \overline{Z} = \dim Z < \prod_{i=1} n_i,$$

where $\overline{Z}$ is the Zariski closure of Z. Next we suppose that $\mathcal{A} \in U := \mathbb{R}^{n_1\times\cdots\times n_k}\setminus\overline{Z}$ and there exist $(V^{(1)}, \ldots, V^{(k)}) \in \mathrm{V}(r-1, n_1) \times \cdots \times \mathrm{V}(r-1, n_k)$ and $(\mu_1, \ldots, \mu_{r-1}) \in \mathbb{R}^{r-1}$ such that $(V^{(1)}, \ldots, V^{(k)}, (\mu_1, \ldots, \mu_{r-1}))$ is a local minimizer of (9). We can write

$$\mathcal{X} := \mathcal{A} - (V^{(1)}, \ldots, V^{(k)}) \cdot \operatorname{diag}(\mu_1, \ldots, \mu_{r-1})$$

as a linear combination of decomposable tensors $\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_k$, such that

(1) $\mathbf{v}_i \in \mathbb{R}^{n_i}$ is a unit norm vector;
(2) either of the following occurs:
  (a) for each $i = 1, \ldots, k$, $\mathbf{v}_i$ is a column vector of $V^{(i)}$;

(b) for each $i = 1, \dots, k$, $\mathbf{v}_i^\intercal V^{(i)} = \mathbf{0}$.

According to the choice of $\mathcal{A}$, there exists $\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_k$ satisfying (1) and (2b) such that

$$\langle \mathcal{X}, \mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_k \rangle \neq 0.$$

Now we set

$$\mathcal{Y} := (V^{(1)}, \dots, V^{(k)}) \cdot \operatorname{diag}(\mu_1, \dots, \mu_{r-1}) + \epsilon \mathbf{v}_1 \otimes \cdots \mathbf{v}_k \in D(\mathbf{n}, r),$$

for a sufficiently small positive number ϵ. We have

$$\|\mathcal{A} - \mathcal{Y}\| = \|\mathcal{X} - \epsilon \mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_k\| < \|\mathcal{X}\|.$$

This contradicts the assumption that $(V^{(1)}, \dots, V^{(k)}, (\mu_1, \dots, \mu_{r-1}))$ is a local minimizer of problem (9). $\square$

As a special case, we suppose that $n_1 = \cdots = n_k = n$ so (56) is written as

$$(r-1)\left(k(n - \frac{r}{2}) + 1\right) < (n - r + 1)^k.$$

We set $r - 1 = (1-\alpha)n$ for $\alpha \in [\frac{1}{n}, 1]$, hence we have

$$(1-\alpha)n\left(\frac{kn(1+\alpha)}{2} + 1 - \frac{k}{2}\right) < \alpha^k n^k.$$

Therefore, to guarantee (56) in this case, it is sufficient to require

$$\frac{2n^{k-2}}{k}\alpha^k + \alpha^2 - 1 > 0. \tag{57}$$

Corollary 5.7. *If $n_1 = \cdots = n_k = n$ and*

$$1 \leqslant r \leqslant \left(1 - \left(\frac{k}{2n^{k-2}}\right)^{\frac{1}{k}}\right) n + 1,$$

then for a generic $\mathcal{A} \in \mathbb{R}^{n \times \cdots \times n}$, any local minimizer of the problem (9) lies in $D(\mathbf{n}, r)$ (or equivalently $U_{\mathbf{n},r}$). In particular, for any fixed k and r, there exists n_0 such that whenever $n \geqslant n_0$ and $\mathcal{A} \in \mathbb{R}^{n \times \cdots \times n}$ is generic, any local minimizer of problem (9) lies in $D(\mathbf{n}, r)$.

Proof. We observe that for any $(k/2n^{k-2})^{1/k} \leqslant \alpha \leqslant 1$, (57) and hence (56) is satisfied by $n_1 = \cdots = n_k = n$ and

$$r = (1-\alpha)n + 1.$$

This implies that for

$$1 \leqslant r \leqslant \left(1 - \left(\frac{k}{2n^{k-2}}\right)^{\frac{1}{k}}\right) n + 1,$$

any local minimizer of the problem (9) lies in $D(\mathbf{n}, r)$ for a generic $\mathcal{A}$. In particular, if $k \geqslant 3$ is a fixed integer, then

$$\lim_{n \to \infty} \frac{\left(1 - \left(\frac{k}{2n^{k-2}}\right)^{\frac{1}{k}}\right) n + 1}{n} = 1.$$

This implies that there exists some integer n_0 such that for a generic $\mathcal{A}$ and any $n \geqslant n_0$, local minimizers of problem (9) must all lie in $D(\mathbf{n}, r)$. $\square$

If (56) is fulfilled, Proposition 5.6 implies that all local maximizers of (10) for a generic tensor are in $D(\mathbf{n}, r)$. Recall from Theorem 3.13 that these local maximizers are finite. Thus, we have the next corollary.

Corollary 5.8. *Let $n_1, \dots, n_k$ and r be positive integers satisfying(56). For a generic tensor $\mathcal{A}$, if Algorithm 4.1 converges to a local maximizer, then the truncation step in Algorithm 4.1 will not be executed when a suitable κ is chosen.*

Combining Corollaries 5.5 and 5.8, we have the following conclusion.

Proposition 5.9. *For almost all LROTA problems, there exist κ, ϵ and τ such that iAPD with Substep 2 being replaced by Algorithm 5.2 reduces to APD after finitely many iterations.*

6. Conclusions

In this paper, we propose an alternating polar decomposition algorithm with adaptive proximal correction and truncation for approximating a given tensor by a low rank orthogonally decomposable tensor. Without any assumption we prove that this algorithm has global convergence and overall sublinear convergence with a sub-optimal explicit convergence rate. For a generic tensor, this algorithm converges R-linearly without any further assumption. For the first time, the convergence rate analysis for the problem of low rank orthogonal tensor approximations is accomplished.

The discussion in Section 5 is all about local maximizers of problem (10). Both APD and iAPD are based on the alternating minimization method [6], which is a variant of gradient ascend. In general, such a method can only converge to a KKT point, including local minimizer, saddle point and local maximizer [5, 6]. However, if each saddle point of a function has the *strict saddle property*, which is guaranteed if it is a nondegenerate critical point, then with probability one, the gradient ascent method converges to a local maximizer [37]. Therefore, for a generic tensor, the proposed algorithm will return a strict local maximizer, since each KKT point is nondegenerate and the objective function is monotonically increasing. More theoretical investigations on this are interesting and important, which will be our next project.

Acknowledgement. This work is partially supported by National Science Foundation of China (Grant No. 11771328). The first author is also partially supported Young Elite Scientists Sponsorship Program by Tianjin and the Natural Science Foundation of Zhejiang Province, China (Grant No. LD19A010002). The second author is also partially supported by National Science Foundation of China (Grant No. 11801548 and Grant No. 11688101), National Key R&D Program of China (Grant No. 2018YFA0306702) and the Recruitment Program of Global Experts of China.

References

[1] P.-A. Absil, R. Mahony, R. Sepulchre, Optimization Algorithms on Matrix Manifolds. Princeton University Press, Princeton, USA, 2008.

[2] A. Anandkumar, R. Ge, D. Hsu, S.M. Kakade, M. Telgarsky, Tensor decompositions for learning latent variable models. J. Mach. Learn. Res., (2014), 15: 2773–2832.

[3] H. Attouch, J. Bolte, P. Redont, A. Soubeyran, Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Łojasiewicz inequality. Math. Oper. Res., (2010), 35: 438–457.

[4] H. Attouch, J. Bolte, B. F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. Math. Program., (2013), 137(1-2): 91–129.

[5] A. Beck, First-Order Methods in Optimization. MOS-SIAM Series on Optimization, SIAM, 2017.

[6] D.P. Bertsekas, Nonlinear Programming. 2nd ed., Athena Scientific, Belmont, USA, 1999.

[7] J. Bochnak, M., Coste, M.-F., Roy, Real Algebraic Geometry. Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 36. Springer, Berlin, 1998.

[8] J. Bolte, A. Daniilidis, O. Ley, L. Mazet, Characterizations of Łojasiewicz inequalities and applications: subgradient flows, talweg, convexity. Trans. Amer. Math. Soc., (2010), 362: 3319–3363.

[9] J. Bolte, S. Sabach, M. Teboulle. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. Math. Program., (2014), 146(1-2): 459–494.

[10] J. Chen, Y. Saad, On the tensor SVD and the optimal low rank orthogonal approximation of tensors. SIAM J. Matrix Anal. Appl., (2009), 30: 1709–1734.

[11] P. Comon, Independent component analysis, a new concept? Signal Process., (1994), 36: 287–314.

[12] P. Comon, Tensors: a brief introduction. IEEE Signal Proc. Mag., (2014), 31: 44–53.

[13] P. Comon, C. Jutten, Handbook of Blind Source Separation. Academic Press, Oxford, 2010.

[14] D. D' Acunto, K. Kurdyka, Explicit bounds for the Łojasiewicz exponent in the gradient inequality for polynomials. Ann. Polon. Math., (2005), 87: 51–61.

[15] L. De Lathauwer, B. De Moor, J. Vandewalle, On the best rank-1 and rank-$(r_1, r_2, , r_N)$ approximation of higher-order tensors. SIAM J. Matrix Anal. Appl., (2000), 21: 1324–1342.

[16] L. De Lathauwer, B. De Moor, J. Vandewalle, Independent component analysis and (simultaneous) third-order tensor diagonalization. IEEE T. Signal Proces., (2001), 49: 2262–2271.

[17] V. De Silva, L.H. Lim, Tensor rank and the ill-posedness of the best low-rank approximation problem. SIAM J. Matrix Anal. Appl., (2008), 29: 1084–1127.

[18] M.P. do Carmo, Riemannian Geometry. Springer, 1992.

[19] C. Eckart, G. Young, The approximation of one matrix by another of lower rank. Psychometrika, (1936), 1: 211–218.

[20] A. Edelman, T.A. Arias, S.T. Smith, The geometry of algorithms with orthogonality constraints. SIAM. J. Matrix Anal. Appl., (1998), 20(2): 303–353.

[21] A. Franc, Etude Algébrique des Multitableaux: Apports de l'Algébre Tensorielle. Thèse de Doctorat, Spécialité Statistiques, Univ. de Montpellier II, Montpellier, France, 1992.

[22] G.H. Golub, C.F. Van Loan, Matrix Computations. 4th ed., Johns Hopkins University Press, Baltimore, MD, 2013.

[23] L. Grasedyck, D. Kressner, C. Tobler, A literature survey of low-rank tensor approximation techniques. GAMM-Mitt., (2013), 36: 53–78.

[24] Y. Guan, D. Chu, Numerical computation for orthogonal low-rank approximation of tensors. SIAM J. Matrix Anal. Appl., (2019), 40: 1047–1065.

[25] W. Hackbusch, Tensor Spaces and Numerical Tensor Calculus. Springer, 2012.

[26] R. Hartshorne, Algebraic Geometry. Graduate Texts in Mathematics 52, Springer, New York, 1977.

[27] N.J. Higham, Computing the polar decompositionwith applications. SIAM J. Sci. Statist. Comput., (1986), 7: 1160–1174.

[28] R.A. Horn, C.R. Johnson, Matrix Analysis. Cambridge University Press, New York, 1985.

[29] S. Hu, G. Li, Convergence rate analysis for the higher order power method in best rank one approximations of tensors. Numer. Math., (2018), 140: 993–1031.

[30] M. Ishteva, P.-A. Absil, P. Van Dooren, Jacobi algorithm for the best low multilinear rank approximation of symmetric tensors. SIAM J. Matrix Anal. Appl., (2013), 34: 651–672.

[31] A.V. Knyazev, M.E. Argentati, Majorization for changes in angles between subspaces, Ritz values, and graph Laplacian spectra. SIAM J. Matrix Anal. Appl., (2006), 29: 15–32.

[32] T.G. Kolda, Orthogonal tensor decompositions. SIAM J. Matrix Anal. Appl., (2001), 23: 243–255.

[33] T.G. Kolda, A counterexample to the possibility of an extension of the Eckart-Young low rank approximation theorem for the orthogonal rank tensor decomposition. SIAM J. Matrix Anal. Appl., (2002), 24: 762–767.

[34] T.G. Kolda, B.W. Bader, Tensor decompositions and applications. SIAM Rev., (2009), 51: 455–500.

[35] J.B. Kruskal, Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. Linear Algebra Appl., (1977), 18: 95–138.
[36] J.M. Landsberg, Tensors: Geometry and Applications. AMS, Providence, RI, 2012.
[37] J.D. Lee, M. Simchowitz, M.I. Jordan, B. Recht, Gradient descent only converges to minimizers. P. Mach. Learn. Res., (2016), 49: 1–12.
[38] G. Li, T.K. Pong, Calculus of the exponent of Kurdyka-Łojasiewicz inequality and its applications to linear convergence of first-order methods. Found. Comput. Math., (2018), 18: 1199-1232.
[39] J. Li, K. Usevich, P. Comon, Globally convergent jacobi-type algorithms for simultaneous orthogonal symmetric tensor diagonalization. SIAM J. Matrix Anal. Appl., (2018), 39: 1–22.
[40] J. Li, K. Usevich, P. Comon, On approximate diagonalization of third order symmetric tensors by orthogonal transformations. Linear Algebra Appl., (2019), 576: 324–351.
[41] J. Li, K. Usevich, P. Comon, Jacobi-type algorithm for low rank orthogonal approximation of symmetric tensors and its convergence analysis. arXiv 1911.00659.
[42] L.-H. Lim, Singular values and eigenvalues of tensors: a variational approach. In Proceedings of the 1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2005, pp. 129–132.
[43] L.-H. Lim, Tensors and hypermatrices. Chapter 15 in Handbook of Linear Algebra, L. Hogben eds, 2013.
[44] H. Liu, W. Wu, A.M.-C. So, Quadratic optimization with orthogonality constraints: explicit Łojasiewicz exponent and linear convergence of line-search methods. Proceedings of the 33rd International Convergnece on Machine Learning, New York, USA, 2016.
[45] S. Łojasiewicz, Une propriété topologique des sous-ensembles analytiques réels, Les Équations aux Dérivées Partielles. Éditions du centre National de la Recherche Scientifique, Paris, 1963, pp. 87–89.
[46] C.D.M. Martin, C.F. Van Loan, A Jacobi-type method for computing orthogonal tensor decompositions. SIAM J. Matrix Anal. Appl., (2008), 30: 1219–1232.
[47] P. McCullagh, Tensor Methods in Statistics. Chapman and Hall, London, (1987)
[48] J. Milnor, Morse Theory. Annals Math. Studies, 51, Princeton Univ. Press, 1963
[49] M.J. Mohlenkamp, Musings on multilinear fitting. Linear Algebra Appl., (2013), 438: 834–852.
[50] C. Mu, D. Hsu, D. Goldfarb, Successive rank-one approximations for nearly orthogonally decomposable symmetric tensors. SIAM J. Matrix Anal. Appl., (2015), 36: 1638–1659.
[51] Y. Nesterov, Introductory Lectures on Convex Optimization-A Basic Course. Kluwer Academic Publishers, 2004.
[52] L. Qi, Z. Luo, Tensor Analysis: Spectral Theory and Special Tensors. SIAM, 2017.
[53] R.T. Rockafellar, R. Wets, Variational Analysis. Grundlehren der Mathematischen Wissenschaften, Vol. 317. Springer, Berlin, 1998.
[54] I.R. Shafarevich, Basic Algebraic Geometry. Springer-Verlag, Berlin, 1977.
[55] M. Sϕrensen, L. De Lathauwer, P. Comon, S. Jcart, L. Deneire, Canonical polyadic decomposition with a columnwise orthonormal factor matrix. SIAM J. Matrix Anal. Appl., (2012), 33: 1190–1213.
[56] J. Sun, C. Chen, Generalized polar decomposition. Math. Numer. Sinica, (1989), 11: 262–273.
[57] A. Uschmajew, Local convergence of the alternating least squares algorithm for canonical tensor approximation. SIAM J. Matrix Anal. Appl., (2012), 33: 639–652.
[58] A. Uschmajew, A new convergence proof for the high-order power method and generalizations. Pacific J. Optim., (2015) 11: 309–321.
[59] K. Usevich, J. Li, P. Comon, Approximate matrix and tensor diagonalization by unitary transformations: convergence of jacobi-type algorithms. arXiv:1905.12295, (2019).
[60] L. Wang, M.T. Chu, On the global convergence of the alternating least squares method for rank-one tensor approximation. SIAM J. Matrix Anal. Appl., (2014), 35(3): 1058–1072.
[61] L. Wang, M.T. Chu, B. Yu, Orthogonal low rank tensor approximation: alternating least squares method and its global convergence. SIAM J. Matrix Anal. Appl., (2015), 36: 1–19.
[62] Y. Yang, The epsilon-alternating least squares for orthogonal low-rank tensor approximation and its global convergence. arXiv 1911.10921.
[63] T. Zhang, G.H. Golub, Rank-one approximation to high order tensors. SIAM J. Matrix Anal. Appl., (2001), 23: 534–550.

Appendix A. Properties on Orthonormal Matrices

A.1. **Polar decomposition.** In this section, we will establish an error bound analysis for the polar decomposition. For a positive semidefinite matrix $H \in \mathrm{S}^n_+$, there exists a unique positive semidefinite matrix $P \in \mathrm{S}^n_+$ such that $P^2 = H$. In the literature, this matrix P is called the *square root* of the matrix H and denoted as $P = \sqrt{H}$. If $H = U\Sigma U^\intercal$ is the eigenvalue decomposition of H, then we have $\sqrt{H} = U\sqrt{\Sigma}U^\intercal$, where $\sqrt{\Sigma}$ is the diagonal matrix whose diagonal elements are square roots of those of Σ. The next result is classical, which can be found in [22, 27].

Lemma A.1 (Polar Decomposition). *Let $A \in \mathbb{R}^{m\times n}$ with $m \geqslant n$. Then there exist an orthonormal matrix $U \in V(n,m)$ and a unique symmetric positive semidefinite matrix $H \in \mathrm{S}^n_+$ such that $A = UH$ and*

$$U \in \operatorname{argmax}\{\langle Q, A\rangle : Q \in V(n,m)\}. \tag{58}$$

Moreover, if A is of full rank, then the matrix U is uniquely determined and H is positive definite.

The matrix decomposition $A = UH$ as in Lemma A.1 is called the *polar decomposition* of the matrix A [22]. For convenience, the matrix U is referred as a *polar orthonormal factor matrix* and the matrix H is the *polar positive semidefinite factor matrix*. The optimization reformulation (58) comes from the approximation problem

$$\min_{Q\in V(n,m)} \|B - QC\|^2$$

for two given matrices B and C of proper sizes. In the following, we give a global error bound for this problem. To this end, the next lemma is useful.

Lemma A.2 (Error Reformulation). *Let p, m, n be positive integers with $m \geqslant n$ and let $B \in \mathbb{R}^{m\times p}$, $C \in \mathbb{R}^{n\times p}$ be two given matrices. We set $A := BC^\intercal \in \mathbb{R}^{m\times n}$ and suppose that $A = WH$ is a polar decomposition of A. We have*

$$\|B - QC\|_F^2 - \|B - WC\|_F^2 = \|W\sqrt{H} - Q\sqrt{H}\|_F^2 \tag{59}$$

for any orthonormal matrix $Q \in V(n,m)$.

Proof. We have

$$\begin{aligned}
\|B - QC\|_F^2 - \|B - WC\|_F^2 &= 2\langle B, WC - QC\rangle \\
&= 2\langle A, W - Q\rangle \\
&= 2\langle WH, W - Q\rangle \\
&= \langle WH, W\rangle - 2\langle WH, Q\rangle + \langle QH, Q\rangle \\
&= \|W\sqrt{H} - Q\sqrt{H}\|_F^2,
\end{aligned}$$

where both the first and the fourth equalities follow from the fact that both Q and W are in $V(n,m)$, and the last one is derived from the fact that H is symmetric and positive semidefinite by Lemma A.1. □

Given an $n \times n$ symmetric positive semidefintie matrix H, we can define a symmetric bilinear form on $\mathbb{R}^{m\times n}$ by

$$\langle P, Q\rangle_H := \langle PH, Q\rangle \tag{60}$$

for all $m \times n$ matrices P and Q. It can also induce a seminorm

$$\|A\|_H := \sqrt{\langle A, A\rangle_H} = \|A\sqrt{H}\|_F. \tag{61}$$

In particular, if H is positive definite, then $\|\cdot\|_H$ is a norm on $\mathbb{R}^{m\times n}$. Thus, the error estimation in (59) can be viewed as a distance estimation between W and Q with respect to the distance induced by this norm. Moreover, if H is the identity matrix, then $\|\cdot\|_H$ is simply the Frobenius norm which induces the Euclidean distance on $\mathbb{R}^{m\times n}$. By Lemma A.2 it is easy to see that the optimizer in (58) is unique whenever A is of full rank.

The following result establishes the error estimation with respect to the Euclidean distance. Given a matrix $A \in \mathbb{R}^{m\times n}$, let $\sigma_{\min}(A)$ be the smallest singular value of A. If A is of full rank, then $\sigma_{\min}(A) > 0$.

Theorem A.3 (Global Error Bound in Frobenius Norm)**.** *Let p, m, n be positive integers with $m \geqslant n$ and let $B \in \mathbb{R}^{m\times p}$ and $C \in \mathbb{R}^{n\times p}$ be two given matrices. We set $A := BC^\intercal \in \mathbb{R}^{m\times n}$ and suppose that A is of full rank with the polar decomposition $A = WH$. We have that*

$$\|B - QC\|_F^2 - \|B - WC\|_F^2 \geqslant \sigma_{\min}(A)\|W - Q\|_F^2 \tag{62}$$

for any orthonormal matrix $Q \in V(n, m)$.

Proof. We know that in this case

$$\sqrt{H} - \sqrt{\sigma_{\min}(A)}I \in \mathrm{S}^n_+ .$$

Therefore we may conclude that

$$\begin{aligned}
\|W\sqrt{H} - Q\sqrt{H}\|_F^2 &= \|W - Q\|^2_{\sqrt{H}} \\
&\geqslant \|W - Q\|^2_{\sqrt{\sigma_{\min}(A)}I} \\
&= \|W\sqrt{\sigma_{\min}(A)}I - Q\sqrt{\sigma_{\min}(A)}I\|_F^2 \\
&= \sigma_{\min}(A)\|W - Q\|_F^2.
\end{aligned}$$

According to Lemma A.2, we can derive the desired inequality. □

Theorem A.3 is a refinement of Sun and Chen's result (cf. [56, Theorem 4.1]), in which the right hand side of (62) has an extra factor $\frac{1}{4}$.

A.2. **Principal angles between subspaces.** Given two linear subspaces $\mathbb{U}$ and $\mathbb{V}$ of dimension r in $\mathbb{R}^n$, the *principal angles* $\{\theta_i : i = 1, \ldots, r\}$ between $\mathbb{U}$ and $\mathbb{V}$ and the associated *principal vectors* $\{(\mathbf{u}_i, \mathbf{v}_i) : i = 1, \ldots, r\}$ are defined recursively by

$$\cos(\theta_i) = \langle \mathbf{u}_i, \mathbf{v}_i\rangle = \max_{\mathbf{u}\in\mathbb{U},\ [\mathbf{u},\mathbf{u}_1,\ldots,\mathbf{u}_{i-1}]\in V(i,n)} \left\{ \max_{\mathbf{v}\in\mathbb{V},\ [\mathbf{v},\mathbf{v}_1,\ldots,\mathbf{v}_{i-1}]\in V(i,n)} \langle \mathbf{u}, \mathbf{v}\rangle \right\}. \tag{63}$$

The following result is standard, whose proof can be found in [22, Section 6.4.3].

Lemma A.4. *For any orthonormal matrices $U, V \in V(r, n)$ of which subspaces spanned by column vectors are $\mathbb{U}$ and $\mathbb{V}$ respectively, we have*

$$\sigma_i(U^\intercal V) = \cos(\theta_i) \text{ for all } i = 1, \ldots, r, \tag{64}$$

where θ_i's are the principal angles between $\mathbb{U}$ and $\mathbb{V}$, and $\sigma_i(U^\intercal V)$ is the i-th largest singular value of the matrix $U^\intercal V$.

Lemma A.5. *For any orthonormal matrices $U, V \in V(r,n)$, we have*

$$\langle U, V\rangle \leqslant \sum_{j=1}^{r} \sigma_j(U^\intercal V). \tag{65}$$

Proof. We recall from [22, pp. 331] that

$$\min_{Q\in \mathrm{O}(r)} \|A - BQ\|_F^2 = \sum_{j=1}^{r} (\sigma_j(A)^2 - 2\sigma_j(B^\intercal A) + \sigma_j(B)^2),$$

for any $n \times r$ matrices A, B. In particular, if $U, V \in \mathrm{V}(r,n)$ then

$$\sigma_j(U) = \sigma_j(V) = 1 \text{ for all } j = 1, \dots, r, \quad \|U\|_F^2 = \|V\|_F^2 = r.$$

This implies that

$$2r - 2\sum_{j=1}^{r} \sigma_j(U^\intercal V) = \min_{Q\in \mathrm{O}(r)} \|U - VQ\|_F^2 \leqslant \|U - V\|_F^2 = 2r - 2\langle U, V\rangle,$$

and the desired inequality follows immediately. □

Lemma A.6. *For any orthonormal matrices $U, V \in V(r,n)$, we have*

$$\|U^\intercal V - I\|_F^2 \leqslant \|U - V\|_F^2. \tag{66}$$

Proof. We have

$$\|U^\intercal V - I\|_F^2 = r + \sum_{i=1}^{r} \cos^2\theta_i - 2\operatorname{tr}(U^\intercal V) \leqslant 2r - 2\operatorname{tr}(U^\intercal V) = \|U - V\|_F^2,$$

where the first equality follows from Lemma A.4. □

Let $\mathbb{U}^\perp$ be the orthogonal complement subspace of a given linear subspace $\mathbb{U}$ in $\mathbb{R}^n$. A useful fact about principal angles between two linear subspaces $\mathbb{U}$ and $\mathbb{V}$ and those between $\mathbb{U}^\perp$ and $\mathbb{V}^\perp$ is stated as follows. The proof can be found in [31, Theorem 2.7].

Lemma A.7. *Let $\mathbb{U}$ and $\mathbb{V}$ be two linear subspaces of the same dimension and let $\frac{\pi}{2} \geqslant \theta_s \geqslant \cdots \geqslant \theta_1 > 0$ be the nonzero principal angles between $\mathbb{U}$ and $\mathbb{V}$. Then the nonzero principal angles between $\mathbb{U}^\perp$ and $\mathbb{V}^\perp$ are $\frac{\pi}{2} \geqslant \theta_s \geqslant \cdots \geqslant \theta_1 > 0$.*

The following result is for the general case, which might be of independent interests.

Lemma A.8. *Let $m \geqslant n$ be positive integers and let $V := [V_1\ V_2] \in \mathrm{O}(m)$ with $V_1 \in V(n,m)$ and $U \in V(n,m)$ be two given orthonormal matrices. Then, there exists an orthonormal matrix $W \in V(m-n,m)$ such that $P := [U\ W] \in \mathrm{O}(m)$ and*

$$\|P - V\|_F^2 \leqslant 2\|U - V_1\|_F^2. \tag{67}$$

Proof. By a simple computation, it is straightforward to verify that (67) is equivalent to

$$\|W - V_2\|_F^2 \leqslant \|U - V_1\|_F^2. \tag{68}$$

To that end, we let $U_2 \in V(m-n,m)$ be an orthonormal matrix such that $[U\ U_2] \in \mathrm{O}(m)$. Then, we have that the linear subspace $\mathbb{U}_2$ spanned by column vectors of U_2 is the orthogonal complement of $\mathbb{U}_1$, which is spanned by column vectors of U. Likewise, let $\mathbb{V}_1$ and $\mathbb{V}_2$ be linear subspaces spanned by column vectors of V_1 and V_2 respectively.

Let $\frac{\pi}{2} \geqslant \theta_s \geqslant \cdots \geqslant \theta_1 > 0$ be the nonzero principal angles between $\mathbb{U}_2$ and $\mathbb{V}_2$ for some nonnegative integer $s \leqslant m-n$. We have by Lemmas A.4, and A.5 that

$$\langle U_2, V_2 \rangle \leqslant \sum_{i=1}^{m-n} \sigma_i(U_2^\intercal V_2) = \sum_{i=1}^{s} \cos(\theta_i) + (m-n) - s. \tag{69}$$

Let $Q \in \mathrm{O}(m-n)$ be a polar orthogonal factor matrix of the matrix $U_2^\intercal V_2$. It follows from the property of polar decomposition that

$$\langle U_2 Q, V_2 \rangle = \sum_{i=1}^{m-n} \sigma_i(U_2^\intercal V_2). \tag{70}$$

On the other hand, nonzero principal angles between $\mathbb{U}_1$ and $\mathbb{V}_1$ are $\frac{\pi}{2} \geqslant \theta_s \geqslant \cdots \geqslant \theta_1 > 0$ by Lemma A.7. Therefore, by Lemmas A.4, and A.5, we have that

$$\langle U, V_1 \rangle \leqslant \sum_{i=1}^{n} \sigma_i(U^\intercal V_1) = \sum_{i=1}^{s} \cos(\theta_i) + n - s. \tag{71}$$

In a conclusion, if we set $W := U_2 Q$, then we have the following:

$$\begin{aligned}
\|W - V_2\|_F^2 &= 2(m-n) - 2\langle U_2 Q, V_2 \rangle \\
&= 2(m-n) - 2\Big(\sum_{i=1}^{s} \cos(\theta_i) + m - n - s\Big) \\
&= 2n - 2\Big(\sum_{i=1}^{s} \cos(\theta_i) + n - s\Big) \\
&\leqslant 2n - 2\langle U, V_1 \rangle \\
&= \|U - V_1\|_F^2,
\end{aligned}$$

where the second equality follows from (69) and (70) and the inequality follows from (71). □

Appendix B. Proofs of Technical Lemmas in Section 4

B.1. Proof of Lemma 4.2.

Proof. For each $i \in \{0, \ldots, k-1\}$, we have

$$\begin{aligned}
f(\mathbb{U}_{i+1,[p]}) - f(\mathbb{U}_{i,[p]}) &= \sum_{j=1}^{r} (\lambda_{j,[p]}^{i+1})^2 - \sum_{j=1}^{r} (\lambda_{j,[p]}^{i})^2 \\
&= \sum_{j=1}^{r} (\lambda_{j,[p]}^{i+1} + \lambda_{j,[p]}^{i})(\lambda_{j,[p]}^{i+1} - \lambda_{j,[p]}^{i}) \\
&= \sum_{j=1}^{r} \lambda_{j,[p]}^{i+1}(\lambda_{j,[p]}^{i+1} - \lambda_{j,[p]}^{i}) + \sum_{j=1}^{r} \lambda_{j,[p]}^{i}(\lambda_{j,[p]}^{i+1} - \lambda_{j,[p]}^{i}).
\end{aligned} \tag{72}$$

We first analyze the second summand in (72) and by considering the following two cases:

(1) If $\sigma_{r,[p]}^{(i+1)} \geqslant \epsilon$, then there is no proximal step in Algorithm 4.1 and we have that $\sigma_{\min}(S_{[p]}^{(i+1)}) = \sigma_{r,[p]}^{(i+1)} \geqslant \epsilon$, where $S_{[p]}^{(i+1)}$ is the polar positive semidefinite factor matrix

of $V^{(i+1)}_{[p]}\Lambda^{(i+1)}_{[p]}$ obtained in (35). From (31), (32) and (33) we notice that

$$\begin{aligned}((U^{(i+1)}_{[p]})^\intercal V^{(i+1)}_{[p]}\Lambda^{(i+1)}_{[p]})_{jj} &= ((\mathbf{u}^{(i+1)}_{j,[p]})^\intercal \mathbf{v}^{(i+1)}_{j,[p]})\lambda^{i}_{j,[p]}\\ &= ((\mathbf{u}^{(i+1)}_{j,[p]})^\intercal \mathcal{A}\tau_{i+1}\mathbf{x}^{(i+1)}_{j,[p]})\lambda^{i}_{j,[p]}\\ &= \mathcal{A}\tau(\mathbf{x}^{(i+2)}_{j,[p]})\lambda^{i}_{j,[p]}\\ &= \lambda^{i+1}_{j,[p]}\lambda^{i}_{j,[p]},\end{aligned}$$

and similarly $((U^{(i+1)}_{[p-1]})^\intercal V^{(i+1)}_{[p]}\Lambda^{(i+1)}_{[p]})_{jj} = \lambda^{i}_{j,[p]}\lambda^{i}_{j,[p]}$. Hence by Lemma A.2, we obtain

$$\begin{aligned}\sum_{j=1}^{r}\lambda^{i}_{j,[p]}(\lambda^{i+1}_{j,[p]}-\lambda^{i}_{j,[p]}) &= \operatorname{Tr}((U^{(i+1)}_{[p]})^\intercal V^{(i+1)}_{[p]}\Lambda^{(i+1)}_{[p]}) - \operatorname{Tr}((U^{(i+1)}_{[p-1]})^\intercal V^{(i+1)}_{[p]}\Lambda^{(i+1)}_{[p]})\\ &= \frac{1}{2}\|(U^{(i+1)}_{[p]}-U^{(i+1)}_{[p-1]})\sqrt{S^{(i+1)}_{[p]}}\|_F^2\\ &\geqslant \frac{\epsilon}{2}\|U^{(i+1)}_{[p]}-U^{(i+1)}_{[p-1]}\|_F^2 \qquad (73)\\ &\geqslant 0.\end{aligned}$$

(2) If $\sigma^{(i+1)}_{r,[p]} < \epsilon$, we consider the following matrix optimization problem

$$\begin{array}{ll}\max & \langle V^{(i+1)}_{[p]}\Lambda^{(i+1)}_{[p]}, U\rangle - \frac{\epsilon}{2}\|U-U^{(i+1)}_{[p-1]}\|_F^2\\ \text{s.t.} & U\in V(r,n_{i+1}).\end{array} \tag{74}$$

Since $U, U^{(i+1)}_{[p-1]}\in V(r,n_{i+1})$, we must have

$$\frac{\epsilon}{2}\|U-U^{(i+1)}_{[p-1]}\|_F^2 = \epsilon r - \epsilon\langle U^{(i+1)}_{[p-1]}, U\rangle.$$

Thus, by Lemma A.1, a global maximizer of (74) is given by a polar orthonormal factor matrix of the matrix $V^{(i+1)}_{[p]}\Lambda^{(i+1)}_{[p]}+\epsilon U^{(i+1)}_{[p-1]}$. By Substep 2 of Algorithm 4.1, $U^{(i+1)}_{[p]}$ is a polar orthonormal factor matrix of the matrix $V^{(i+1)}_{[p]}\Lambda^{(i+1)}_{[p]}+\epsilon U^{(i+1)}_{[p-1]}$, and hence a global maximizer of (74). Thus, by the optimality of $U^{(i+1)}_{[p]}$ for (74), we have

$$\langle V^{(i+1)}_{[p]}\Lambda^{(i+1)}_{[p]}, U^{(i+1)}_{[p]}\rangle - \frac{\epsilon}{2}\|U^{(i+1)}_{[p]}-U^{(i+1)}_{[p-1]}\|_F^2 \geqslant \langle V^{(i+1)}_{[p]}\Lambda^{(i+1)}_{[p]}, U^{(i+1)}_{[p-1]}\rangle.$$

Therefore, the inequality (73) in case (1) also holds in this case.

Consequently, we have

$$0\leqslant \sum_{j=1}^{r}\lambda^{i}_{j,[p]}(\lambda^{i+1}_{j,[p]}-\lambda^{i}_{j,[p]}) = \sum_{j=1}^{r}\lambda^{i+1}_{j,[p]}\lambda^{i}_{j,[p]} - \sum_{j=1}^{r}(\lambda^{i}_{j,[p]})^2, \tag{75}$$

which together with Cauchy-Schwartz inequality implies that

$$\Big(\sum_{j=1}^{r}(\lambda^{i}_{j,[p]})^2\Big)^2 \leqslant \Big(\sum_{j=1}^{r}\lambda^{i}_{j,[p]}\lambda^{i+1}_{j,[p]}\Big)^2 \leqslant \sum_{j=1}^{r}(\lambda^{i}_{j,[p]})^2\sum_{j=1}^{r}(\lambda^{i+1}_{j,[p]})^2. \tag{76}$$

Since $f(\mathbb{U}_{[0]}) > 0$, we conclude that $f(\mathbb{U}_{0,[1]}) = \sum_{j=1}^{r}(\lambda_{j,[1]}^{0})^2 > 0$ and hence $\sum_{j=1}^{r} \lambda_{j,[1]}^{1}\lambda_{j,[1]}^{0} > 0$ by (75). Thus, we conclude that

$$\sum_{j=1}^{r} \lambda_{j,[1]}^{1}\lambda_{j,[1]}^{0} \leqslant \sum_{j=1}^{r}(\lambda_{j,[1]}^{1})^2 \frac{\sum_{j=1}^{r}(\lambda_{j,[1]}^{0})^2}{\sum_{j=1}^{r}\lambda_{j,[1]}^{1}\lambda_{j,[1]}^{0}} \leqslant \sum_{j=1}^{r}(\lambda_{j,[1]}^{1})^2, \tag{77}$$

where the first inequality follows from (76) and the second from (75). Combining (77) with (72) and (73), we may obtain (40) for $i = 0$ and $p = 1$.

On the one hand, from (76) we obtain

$$0 \leqslant \sum_{j=1}^{r}(\lambda_{j,[p]}^{i})^2 \Big(\sum_{j=1}^{r}(\lambda_{j,[p]}^{i+1})^2 - \sum_{j=1}^{r}(\lambda_{j,[p]}^{i})^2\Big).$$

Since $\sum_{j=1}^{r}(\lambda_{j,[p]}^{i})^2 > 0$ if there is no truncation, we must have

$$0 \leqslant \sum_{j=1}^{r}(\lambda_{j,[p]}^{i+1})^2 - \sum_{j=1}^{r}(\lambda_{j,[p]}^{i})^2 = f(\mathbb{U}_{i+1,[p]}) - f(\mathbb{U}_{i,[p]}),$$

i.e., the objective function f is monotonically increasing during the APD iteration as long as there is no truncation. On the other hand, there are at most r truncations and hence the total loss of f by the truncation is at most $r\kappa^2 < f(\mathbb{U}_{[0]})$. Therefore, f is always positive and according to (75), we may conclude that $\sum_{j=1}^{r} \lambda_{j,[p]}^{i+1}\lambda_{j,[p]}^{i} > 0$ along iterations. By induction on p, we obtain

$$\sum_{j=1}^{r} \lambda_{j,[p]}^{i+1}\lambda_{j,[p]}^{i} \leqslant \sum_{j=1}^{r}(\lambda_{j,[p]}^{i+1})^2 \frac{\sum_{j=1}^{r}(\lambda_{j,[p]}^{i})^2}{\sum_{j=1}^{r}\lambda_{j,[p]}^{i+1}\lambda_{j,[p]}^{i}} \leqslant \sum_{j=1}^{r}(\lambda_{j,[p]}^{i+1})^2,$$

which together with (72) and (73), implies (40) for arbitrary nonnegative integer p. $\square$

B.2. **Proof for Lemma 4.5.**

Proof. The subdifferentials of h can be partitioned as follows:

$$\partial h(\mathbb{U}) = (\nabla_1 f(\mathbb{U}) + \partial\delta_{V(r,n_1)}(U^{(1)})) \times \cdots \times (\nabla_k f(\mathbb{U}) + \partial\delta_{V(r,n_k)}(U^{(k)})). \tag{78}$$

Following the notation of Algorithm 4.1, we set

$$\mathbf{x}_j := (\mathbf{u}_{j,[p+1]}^{(1)}, \ldots, \mathbf{u}_{j,[p+1]}^{(k)}) \text{ for all } j = 1, \ldots, r,$$

where $\mathbf{u}_{j,[p+1]}^{(i)}$ is the j-th column of the matrix $U_{[p+1]}^{(i)}$ for all $i \in \{1, \ldots, k\}$,

$$V^{(i)} := \begin{bmatrix} \mathbf{v}_1^{(i)} & \ldots & \mathbf{v}_r^{(i)} \end{bmatrix} \text{ with } \mathbf{v}_j^{(i)} := \mathcal{A}\tau_i(\mathbf{x}_j) \text{ for all } j = 1, \ldots, r, \tag{79}$$

for all $i \in \{1, \ldots, k\}$ and

$$\Lambda := \operatorname{diag}(\lambda_1, \ldots, \lambda_r) \text{ with } \lambda_j := \mathcal{A}\tau(\mathbf{x}_j). \tag{80}$$

By (35) and (74), we have

$$V_{[p+1]}^{(i)}\Lambda_{[p+1]}^{(i)} + \alpha U_{[p]}^{(i)} = U_{[p+1]}^{(i)}S_{[p+1]}^{(i)}, \tag{81}$$

where $\alpha \in \{0, \epsilon\}$ depending on whether or not there is a proximal correction. According to (3) and (81), we have

$$-U_{[p+1]}^{(i)} \in \partial\delta_{V(r,n_i)}(U_{[p+1]}^{(i)}), \quad V_{[p+1]}^{(i)}\Lambda_{[p+1]}^{(i)} + \alpha\big(U_{[p]}^{(i)}\big) \in \partial\delta_{V(r,n_i)}(U_{[p+1]}^{(i)}),$$

which implies that $V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]}+\alpha\big(U^{(i)}_{[p]}-U^{(i)}_{[p+1]}\big)\in\partial\delta_{V(r,n_i)}(U^{(i)}_{[p+1]})$. If we take

$$W^{(i)}_{[p+1]}:=2V^{(i)}\Lambda-2V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]}-2\alpha\big(U^{(i)}_{[p]}-U^{(i)}_{[p+1]}\big), \tag{82}$$

then we have

$$W^{(i)}_{[p+1]}\in 2V^{(i)}\Lambda+\partial\delta_{V(r,n_i)}(U^{(i)}_{[p+1]}).$$

On the other hand,

$$\begin{aligned}
&\frac{1}{2}\|W^{(i)}_{[p+1]}\|_F\\
&=\|V^{(i)}\Lambda-V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]}-\alpha\big(U^{(i)}_{[p]}-U^{(i)}_{[p+1]}\big)\|_F\\
&\leqslant\|V^{(i)}\Lambda-V^{(i)}_{[p+1]}\Lambda\|_F+\|V^{(i)}_{[p+1]}\Lambda-V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]}\|_F+\alpha\|U^{(i)}_{[p]}-U^{(i)}_{[p+1]}\|_F\\
&\leqslant\|V^{(i)}-V^{(i)}_{[p+1]}\|_F\|\Lambda\|_F+\|V^{(i)}_{[p+1]}\|_F\|\Lambda-\Lambda^{(i)}_{[p+1]}\|_F+\alpha\|U^{(i)}_{[p]}-U^{(i)}_{[p+1]}\|_F\\
&\leqslant\|\mathcal{A}\|\|\Lambda\|_F\big(\sum_{j=1}^{r}\|\tau_i(\mathbf{x}_j)-\tau_i(\mathbf{x}^i_{j,[p+1]})\|\big)\\
&\quad+\|V^{(i)}_{[p+1]}\|_F\|\mathcal{A}\|\big(\sum_{j=1}^{r}\|\tau(\mathbf{x}_j)-\tau(\mathbf{x}^i_{j,[p+1]})\|\big)+\alpha\|U^{(i)}_{[p]}-U^{(i)}_{[p+1]}\|_F\\
&\leqslant\sqrt{r}\|\mathcal{A}\|^2\big(\sum_{j=1}^{r}\sum_{s=i+1}^{k}\|\mathbf{u}^{(s)}_{j,[p+1]}-\mathbf{u}^{(s)}_{j,[p]}\|\big)\\
&\quad+\sqrt{r}\|\mathcal{A}\|^2\big(\sum_{j=1}^{r}\sum_{s=i}^{k}\|\mathbf{u}^{(s)}_{j,[p+1]}-\mathbf{u}^{(s)}_{j,[p]}\|\big)+\alpha\|U^{(i)}_{[p]}-U^{(i)}_{[p+1]}\|_F\\
&\leqslant(2r\sqrt{r}\|\mathcal{A}\|^2+\epsilon)\|\mathbb{U}_{[p+1]}-\mathbb{U}_{[p]}\|_F,
\end{aligned}$$

where the third inequality follows from the fact that

$$V^{(i)}-V^{(i)}_{[p+1]}=\begin{bmatrix}\mathcal{A}(\tau_i(\mathbf{x}_1)-\tau_i(\mathbf{x}^i_{1,[p+1]})) & \dots & \mathcal{A}(\tau_i(\mathbf{x}_r)-\tau_i(\mathbf{x}^i_{r,[p+1]}))\end{bmatrix},$$

and a similar formula for $\Lambda-\Lambda^{(i)}_{[p+1]}$, the fourth follows from the fact that

$$|\mathcal{A}\tau(\mathbf{x})|\leqslant\|\mathcal{A}\|$$

for any vector $\mathbf{x}:=(\mathbf{x}_1,\dots,\mathbf{x}_k)$ with $\|\mathbf{x}_i\|=1$ for all $i=1,\dots,k$ and the last one follows from $\alpha\leqslant\epsilon$. This, together with (78), implies (43). □

B.3. **Proof for Lemma 4.10.**

Proof. We let

$$W^{(i)}_{[p+1]}:=V^{(i)}\Lambda-V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]}-\alpha\big(U^{(i)}_{[p]}-U^{(i)}_{[p+1]}\big),$$

where $\alpha\in\{0,\epsilon\}$ depending on whether there is a proximal correction or not (cf. the proof for Lemma 4.5). It follows from Lemma 4.5 that

$$\|W^{(i)}_{[p+1]}\|_F\leqslant\gamma_0\|U^{(i)}_{[p]}-U^{(i)}_{[p+1]}\|_F$$

for some constant $\gamma_0>0$. By Algorithm 4.1, we have

$$V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]}+\alpha U^{(i)}_{[p]}=U^{(i)}_{[p+1]}S^{(i)}_{[p+1]} \tag{83}$$

where $S^{(i)}_{[p+1]}$ is a symmetric positive semidefinite matrix. Since $U^{(i)}_{[p+1]} \in V(r, n_i)$ is an orthonormal matrix, we have

$$S^{(i)}_{[p+1]} = (U^{(i)}_{[p+1]})^\intercal \big(V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]} + \alpha U^{(i)}_{[p]}\big) = \big(V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]} + \alpha U^{(i)}_{[p]}\big)^\intercal U^{(i)}_{[p+1]}, \tag{84}$$

where the second equality follows from the symmetry of the matrix $S^{(i)}_{[p+1]}$.

Consequently, we have

$$\begin{aligned}
&\frac{1}{2}\|\nabla_i f(\mathbb{U}_{[p+1]}) - U^{(i)}_{[p+1]}(\nabla_i f(\mathbb{U}_{[p+1]}))^\intercal U^{(i)}_{[p+1]}\|_F \\
&= \|V^{(i)}\Lambda - U^{(i)}_{[p+1]}(V^{(i)}\Lambda)^\intercal U^{(i)}_{[p+1]}\|_F \\
&= \|W^{(i)}_{[p+1]} + V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]} + \alpha\big(U^{(i)}_{[p]} - U^{(i)}_{[p+1]}\big) - U^{(i)}_{[p+1]}(V^{(i)}\Lambda)^\intercal U^{(i)}_{[p+1]}\|_F \\
&\leqslant \|W^{(i)}_{[p+1]}\|_F + \alpha\|U^{(i)}_{[p]} - U^{(i)}_{[p+1]}\|_F + \|V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]} - U^{(i)}_{[p+1]}(V^{(i)}\Lambda)^\intercal U^{(i)}_{[p+1]}\|_F \\
&\leqslant \gamma_1\|U^{(i)}_{[p]} - U^{(i)}_{[p+1]}\|_F + \|V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]} - U^{(i)}_{[p+1]}(V^{(i)}\Lambda)^\intercal U^{(i)}_{[p+1]}\|_F,
\end{aligned} \tag{85}$$

where $\gamma_1 = \gamma_0 + \epsilon$. Next, we derive an estimation for the second summand of the right hand side of (85). To do this, we notice that

$$\begin{aligned}
&\|V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]} - U^{(i)}_{[p+1]}(V^{(i)}\Lambda)^\intercal U^{(i)}_{[p+1]}\|_F \\
&= \|U^{(i)}_{[p+1]}S^{(i)}_{[p+1]} - \alpha U^{(i)}_{[p]} - U^{(i)}_{[p+1]}(V^{(i)}\Lambda)^\intercal U^{(i)}_{[p+1]}\|_F \\
&= \|U^{(i)}_{[p+1]}\big(V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]} + \alpha U^{(i)}_{[p]}\big)^\intercal U^{(i)}_{[p+1]} - \alpha U^{(i)}_{[p]} - U^{(i)}_{[p+1]}(V^{(i)}\Lambda)^\intercal U^{(i)}_{[p+1]}\|_F \\
&\leqslant \|U^{(i)}_{[p+1]}\big((V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]})^\intercal - (V^{(i)}\Lambda)^\intercal\big)U^{(i)}_{[p+1]}\|_F + \alpha\|U^{(i)}_{[p+1]}(U^{(i)}_{[p]})^\intercal U^{(i)}_{[p+1]} - U^{(i)}_{[p]}\|_F \\
&\leqslant \|V^{(i)}_{[p+1]}\Lambda^{(i)}_{[p+1]} - V^{(i)}\Lambda\|_F + \alpha\|U^{(i)}_{[p+1]}(U^{(i)}_{[p]})^\intercal U^{(i)}_{[p+1]} - U^{(i)}_{[p]}\|_F \\
&\leqslant \gamma_2\|U^{(i)}_{[p+1]} - U^{(i)}_{[p]}\|_F,
\end{aligned} \tag{86}$$

where $\gamma_2 > 0$ is some constant, the first equality follows from (83), the second from (84), the second inequality[2] follows from the fact that $U^{(i)}_{[p+1]} \in V(r, n)$ and the last inequality from the relation

$$\|(U^{(i)}_{[p]})^\intercal U^{(i)}_{[p+1]} - I\|_F \leqslant \|U^{(i)}_{[p]} - U^{(i)}_{[p+1]}\|_F,$$

which is obtained by Lemma A.6. The desired inequality can be derived easily from (85) and (86).

□

Department of Mathematics, School of Science, Hangzhou Dianzi University, Hangzhou 310018, China.
E-mail address: shenglonghu@hdu.edu.cn

KLMM, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China.
E-mail address: keyk@amss.ac.cn

[2]Since U is orthonormal, we must have

$$\|UAU\|_F^2 = \|AU\|_F^2 = \langle AUU^\intercal, A\rangle \leqslant \|A\|_F^2.$$