



HAL
open science

On convergence of matrix and tensor approximate diagonalization algorithms by unitary transformations

Konstantin Usevich, Jianze Li, Pierre Comon

► **To cite this version:**

Konstantin Usevich, Jianze Li, Pierre Comon. On convergence of matrix and tensor approximate diagonalization algorithms by unitary transformations. 2019. hal-01998900v1

HAL Id: hal-01998900

<https://hal.science/hal-01998900v1>

Preprint submitted on 29 Jan 2019 (v1), last revised 5 Jul 2020 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On convergence of matrix and tensor approximate diagonalization algorithms by unitary transformations

Konstantin Usevich, Jianze Li, Pierre Comon

January 29, 2019

Abstract

In this paper, we propose a gradient-based Jacobi algorithm for approximate diagonalization of complex matrices and tensors by unitary transformations. We provide global and local convergence results for this algorithm. The convergence results also apply to the case of real-valued tensors.

1 Introduction

In this paper, we consider the following optimization problem

$$\mathbf{U}_* = \arg \max_{\mathbf{U} \in \mathcal{U}_n} f(\mathbf{U}), \quad (1)$$

where $\mathcal{U}_n \subseteq \mathbb{C}^{n \times n}$ is the unitary group and

$$f : \mathcal{U}_n \rightarrow \mathbb{R}^+ \quad (2)$$

is a differentiable function. In particular, we focus on several cost functions that are motivated by blind source separation problems. Jacobi-type algorithms are widely used for maximization of these cost functions, due to the simplicity of Jacobi rotations. Nevertheless, the convergence of these algorithms was not well-studied so far.

In the real-valued case (for the orthogonal group), a gradient-based Jacobi-type algorithm (which we call *Jacobi-G*) was proposed in [22], and the global/local convergence of this algorithm was proved [22, 25]. In this paper, we show that it is possible to extend the Jacobi-G algorithm to the case of \mathcal{U}_n for any differentiable function (2), and prove convergence properties similar to those in [22, 25] under some conditions.

The main contributions of the paper are:

- We generalize the Jacobi-G algorithm to the complex case, namely to the unitary group, and prove its convergence to stationary points and global rates of convergence.
- For the case of matrix and tensor diagonalization, we obtain local linear convergence to local maxima satisfying second order regularity conditions. These results also apply to the real-valued case (orthogonal diagonalization) considered in [25].

i.e., the matrix with the same elements as \mathbf{I}_n except that

$$\mathcal{P}_{i,j}(\mathbf{G}^{(i,j,\Psi)}) = \Psi.$$

It is obvious that the set of matrices $\mathbf{G}^{(i,j,\Psi)}$ is canonically isomorphic to \mathcal{U}_2 , and it defines a subgroup of \mathcal{U}_n . Now, for a fixed unitary matrix $\mathbf{U} \in \mathcal{U}_n$, we define the restriction of (2) to rotations in the (i, j) -th plane as:

$$\begin{aligned} h_{(i,j),\mathbf{U}} : \mathcal{U}_2 &\longrightarrow \mathbb{R}^+ \\ \Psi &\longmapsto f(\mathbf{U}\mathbf{G}^{(i,j,\Psi)}). \end{aligned} \quad (4)$$

The classic Jacobi-C algorithm can be formulated as follows.

Algorithm 1. (*General Jacobi-C algorithm*)

Input: A differentiable function $f : \mathcal{U}_n \rightarrow \mathbb{R}^+$ defined in a neighborhood of \mathcal{U}_n , a starting point \mathbf{U}_0 .

Output: Sequence of iterations \mathbf{U}_k .

- **For** $k = 1, 2, \dots$ until a stopping criterion is satisfied do
- Choose an index pair (i_k, j_k) in the following cyclic order:

$$\begin{aligned} (1, 2) &\rightarrow (1, 3) \rightarrow \dots \rightarrow (1, n) \rightarrow \\ (2, 3) &\rightarrow \dots \rightarrow (2, n) \rightarrow \\ \dots &\rightarrow (n-1, n) \rightarrow \\ (1, 2) &\rightarrow (1, 3) \rightarrow \dots \end{aligned} \quad (5)$$

- Find Ψ_k that maximizes $h_k(\Psi) \stackrel{\text{def}}{=} h_{(i_k, j_k), \mathbf{U}_{k-1}}(\Psi)$.
- Update $\mathbf{U}_k = \mathbf{U}_{k-1} \mathbf{G}^{(i_k, j_k, \Psi_k)}$.
- **End for**

2.2 Cost functions under consideration

In this paper, we mainly consider the following three cost functions from \mathcal{U}_n to \mathbb{R}^+ .

- (i) *Approximate diagonalization of a set of matrices.* Let $\mathbf{A}^{(\ell)} = \mathbf{A}^{(\ell, \Re)} + i\mathbf{A}^{(\ell, \Im)} \in \mathbb{C}^{n \times n}$, $1 \leq \ell \leq L$, be a set of Hermitian¹ matrices. The cost function is defined as

$$f(\mathbf{U}) = \sum_{\ell=1}^L \|\text{diag}\{\mathbf{W}^{(\ell)}\}\|^2, \quad (6)$$

where $\mathbf{W}^{(\ell)} = \mathbf{U}^H \mathbf{A}^{(\ell)} \mathbf{U}$.

¹In fact, it is easy to consider also non-Hermitian matrices, as shown in [12].

- (ii) *Approximate diagonalization of a partially symmetric 3rd order tensor.* Let $\mathcal{A} \in \mathbb{C}^{n \times n \times n}$ be a tensor satisfying the partial symmetry condition:

$$\mathcal{A}_{ijk} = \mathcal{A}_{ikj} \quad (7)$$

for any $1 \leq i, j, k \leq n$. The cost function is defined as:

$$f(\mathbf{U}) = \|\text{diag}\{\mathcal{W}\}\|^2, \quad (8)$$

where $\mathcal{W}_{ijk} = \sum_{p,q,r} \mathcal{A}_{pqr} U_{pi}^* U_{qj} U_{rk}$.

- (iii) *Approximate diagonalization of a partially symmetric 4th order tensor.* Let $\mathcal{B} \in \mathbb{C}^{n \times n \times n \times n}$ be a tensor satisfying

$$\mathcal{B}_{ijkl} = \mathcal{B}_{jilk} \quad \text{and} \quad \mathcal{B}_{ijkl} = \mathcal{B}_{klij}^* \quad (9)$$

for any $1 \leq i, j, k, l \leq n$. The cost function is defined as

$$f(\mathbf{U}) = \sum_{p=1}^n \mathcal{V}_{pppp}, \quad (10)$$

where $\mathcal{V}_{ijkl} = \sum_{p,q,r,s} \mathcal{B}_{pqrs} U_{pi}^* U_{qj}^* U_{rk} U_{sl}$.

Remark 2.1. *As in (6), the simultaneous diagonalization problem can be also considered for 3rd and 4th order tensors. In this paper however, we prefer to consider the single tensor case in (8) and (10) (for simplicity).*

The motivation behind these cost functions comes from blind source separation.

- (i) The cost function (6) is used for joint diagonalization of covariance matrices [11, 12].
- (ii) An example of the 3rd order tensor satisfying property (7) is the cumulant tensor with $\mathcal{A}_{ijk} = \text{Cum}(v_i, v_j^*, v_k^*)$, where \mathbf{v} is a complex random vector [16].
- (iii) An example of the 4th order tensor satisfying property (9) is the cumulant tensor with $\mathcal{A}_{ijkl} = \text{Cum}(v_i, v_j, v_k^*, v_l^*)$, where \mathbf{v} is a complex random vector [13].

Note that many other differentiable functions may be used in this framework [14].

Remark 2.2. *For simplicity, we consider in this paper only fully contravariant tensors [27].*

2.3 Elementary rotations

In this paper, we assume that the cost function (1) satisfies that $f(\mathbf{U})$ is invariant under permutations of columns of \mathbf{U} and multiplications of columns of \mathbf{U} by complex scalars of modulus 1. It can be seen that the cost functions in Section 2.2 satisfy this condition. In this case, we see that the restriction (4) satisfies

$$h_{(i,j),\mathbf{U}}(\Psi) = h_{(i,j),\mathbf{U}}\left(\Psi \begin{bmatrix} z_1 & 0 \\ 0 & z_2 \end{bmatrix}\right) \quad (11)$$

for any $|z_1| = |z_2| = 1$. Hence, to maximize $h_k(\Psi)$ in Algorithm 1, we can set

$$\begin{aligned}\Psi &= \Psi(c, s_1, s_2) = \begin{bmatrix} c & -s \\ s^* & c \end{bmatrix} = \begin{bmatrix} c & -(s_1 + is_2) \\ s_1 - is_2 & c \end{bmatrix} \\ &= \begin{bmatrix} \cos \theta & -\sin \theta e^{i\phi} \\ \sin \theta e^{-i\phi} & \cos \theta \end{bmatrix} = \frac{1}{\sqrt{1+|z|^2}} \begin{bmatrix} 1 & -z \\ z^* & 1 \end{bmatrix} \in \mathcal{SU}_2, \end{aligned} \quad (12)$$

with $c \in \mathbb{R}^+$ and $s = s_1 + is_2 = \sin \theta e^{i\phi} \in \mathbb{C}$ satisfying $c^2 + |s|^2 = 1$. In this sense, we will often use the notations

$$h(c, s_1, s_2) = h_{(i,j),\mathbf{U}}(c, s_1, s_2) = h_{(i,j),\mathbf{U}} \left(\begin{bmatrix} c & -s \\ s^* & c \end{bmatrix} \right),$$

alternatively. Also denote

$$\mathbf{r} = \mathbf{r}(c, s_1, s_2) \stackrel{\text{def}}{=} (\cos 2\theta, -\sin 2\theta \sin \phi, -\sin 2\theta \cos \phi)^\top = (2c^2 - 1, -2cs_1, -2cs_2)^\top. \quad (13)$$

Now, we show that, for the cost functions in Section 2.2, the Jacobi rotations can be computed in a unified way, by finding an eigenvector of a 3×3 real matrix.

Lemma 2.3. *Let $\mathbf{U} \in \mathcal{U}_n$. Then all the cost functions in Section 2.2 have the form*

$$h_{(i,j),\mathbf{U}}(c, s_1, s_2) = \mathbf{r}^\top \mathbf{\Gamma}^{(i,j),\mathbf{U}} \mathbf{r} + C, \quad (14)$$

where the constant C does not depend on (c, s_1, s_2) , and $\mathbf{\Gamma}^{(i,j),\mathbf{U}}$ is a symmetric matrix in $\mathbb{R}^{3 \times 3}$ defined as follows:

(i) for the cost function (6), we have $C = \frac{1}{2} \sum_{\ell=1}^L (W_{jj}^{(\ell)} + W_{ii}^{(\ell)})^2$ and

$$\begin{aligned}\mathbf{\Gamma}^{(i,j),\mathbf{U}} &= \sum_{\ell=1}^L \begin{bmatrix} \frac{1}{2}(W_{jj}^{(\ell)} - W_{ii}^{(\ell)})^2 & (W_{jj}^{(\ell)} - W_{ii}^{(\ell)})W_{ij}^{(\ell,\Re)} & (W_{jj}^{(\ell)} - W_{ii}^{(\ell)})W_{ij}^{(\ell,\Im)} \\ (W_{jj}^{(\ell)} - W_{ii}^{(\ell)})W_{ij}^{(\ell,\Re)} & 2(W_{ij}^{(\ell,\Re)})^2 & 2W_{ij}^{(\ell,\Re)}W_{ij}^{(\ell,\Im)} \\ (W_{jj}^{(\ell)} - W_{ii}^{(\ell)})W_{ij}^{(\ell,\Im)} & 2W_{ij}^{(\ell,\Re)}W_{ij}^{(\ell,\Im)} & 2(W_{ij}^{(\ell,\Im)})^2 \end{bmatrix} \\ &= \frac{1}{2} \sum_{\ell=1}^L \begin{bmatrix} (W_{jj}^{(\ell)} - W_{ii}^{(\ell)}) \\ 2W_{ij}^{(\ell,\Re)} \\ 2W_{ij}^{(\ell,\Im)} \end{bmatrix} \begin{bmatrix} (W_{jj}^{(\ell)} - W_{ii}^{(\ell)}) & 2W_{ij}^{(\ell,\Re)} & 2W_{ij}^{(\ell,\Im)} \end{bmatrix};\end{aligned}$$

(ii) for the cost function (8), we have $C = 0$ and $\mathbf{\Gamma} = \mathbf{\Gamma}^{(i,j),\mathbf{U}}$ (see [16, (9.29)] and [15, Section 5.3.2]) is:

$$\begin{aligned}\Gamma_{11} &= a_1, \quad \Gamma_{12} = \Im(v_1) + \Im(v_2), \quad \Gamma_{13} = \Re(v_1) - \Re(v_2), \quad \Gamma_{22} = v_4 - \Re(v_3), \\ \Gamma_{23} &= \Im(v_3), \quad \Gamma_{33} = v_4 + \Re(v_3), \\ a_1 &= |\mathcal{W}_{111}|^2 + |\mathcal{W}_{222}|^2, \quad a_2 = |\mathcal{W}_{112}|^2 + |\mathcal{W}_{212}|^2, \quad a_3 = |\mathcal{W}_{211}|^2 + |\mathcal{W}_{122}|^2, \\ a_4 &= \mathcal{W}_{111}\mathcal{W}_{112}, \quad a_5 = \mathcal{W}_{111}\mathcal{W}_{211}, \quad a_6 = \mathcal{W}_{222}\mathcal{W}_{122}, \quad a_7 = \mathcal{W}_{222}\mathcal{W}_{212}, \\ a_8 &= \mathcal{W}_{111}\mathcal{W}_{212} + \mathcal{W}_{222}\mathcal{W}_{112}, \quad a_9 = \mathcal{W}_{111}\mathcal{W}_{122} + \mathcal{W}_{222}\mathcal{W}_{211}, \\ a_{10} &= \mathcal{W}_{211}\mathcal{W}_{112} + \mathcal{W}_{122}\mathcal{W}_{212}, \\ v_1 &= a_7 - a_5/2, \quad v_2 = a_4 - a_6/2, \quad v_3 = a_9/2 + a_{10}, \\ v_4 &= (a_1 + a_3)/4 + a_2 + \Re(a_8).\end{aligned}$$

(iii) for the cost function (10), we have $C = 0$ and $\mathbf{\Gamma} = \mathbf{\Gamma}^{(i,j,U)}$ [13] is

$$\begin{aligned}\Gamma_{11} &= \mathcal{V}_{iiii} + \mathcal{V}_{jjjj}, & \Gamma_{12} &= \mathcal{V}_{ijjj}^{\Re} - \mathcal{V}_{iiij}^{\Re}, & \Gamma_{13} &= \mathcal{V}_{ijjj}^{\Im} - \mathcal{V}_{iiij}^{\Im}, & \Gamma_{23} &= \mathcal{V}_{iiij}^{\Im}, \\ \Gamma_{22} &= \frac{1}{2}(\mathcal{V}_{iiii} + \mathcal{V}_{jjjj}) + 2\mathcal{V}_{ijij} + \mathcal{V}_{iiij}^{\Re}, & \Gamma_{33} &= \frac{1}{2}(\mathcal{V}_{iiii} + \mathcal{V}_{jjjj}) + 2\mathcal{V}_{ijij} - \mathcal{V}_{iiij}^{\Re}.\end{aligned}$$

Remark 2.4. (i) By Lemma 2.3, the maximization of $h_{(i,j,U)}(c, s_1, s_2)$ is equivalent to maximization of the quadratic form $\mathbf{r}^\top \mathbf{\Gamma}^{(i,j,U)} \mathbf{r}$ on the unit sphere $\|\mathbf{r}\| = 1$. Thus, in the iteration of Algorithm 2, \mathbf{r} can be taken as the eigenvector corresponding to the maximal eigenvalue of $\mathbf{\Gamma}^{(i,j,U)}$ (or an eigenvector if the two leading eigenvalues are equal), in which case, we denote this eigenvector by \mathbf{w} .

(ii) Even if the eigenvector is unique, it is defined up to a sign change. We choose the sign such that $w_1 = \cos 2\theta = 2c^2 - 1 \geq 0$ in (13). Hence, we can guarantee that $\theta \in [0, \frac{\pi}{4}]$ and $c > \frac{\sqrt{2}}{2}$. More precisely, we set $\theta = \arccos(w_1)/2 \in [0, \frac{\pi}{4}]$ and $s_1 = -\frac{w_2}{2\cos\theta}$, $s_2 = -\frac{w_3}{2\cos\theta}$.

3 Jacobi-G algorithm for unitary matrices \mathcal{U}_n

Before we formulate the Jacobi-G algorithm, it is necessary to recall some facts about derivatives of functions defined on manifolds and complex derivatives.

3.1 Riemannian gradient for embedded submanifolds of \mathbb{R}^N

We recall some basics on real manifolds that are submanifolds of an Euclidean space (the reader is invited to consult [6] if needed). Consider an m -dimensional submanifold $\mathcal{M} \subseteq \mathcal{E}$ of the N -dimensional Euclidean space $\mathcal{E} \simeq \mathbb{R}^N$ (in what follows we will assume $\mathcal{E} = \mathbb{R}^N$ for simplicity). Then the tangent space on the manifold can be associated with an m -dimensional subspace $\mathbf{T}_x \mathcal{M} \subseteq \mathcal{E}$. We will assume that $\mathbf{T}_x \mathcal{M}$ is equipped with the Riemannian metric induced by the ambient space, i.e.,

$$\langle \eta, \xi \rangle = \langle \eta, \xi \rangle_{\mathcal{E}},$$

i.e., it is just the inner product of vectors in \mathcal{E} .

Consider the function $f : \mathcal{M} \rightarrow \mathbb{R}$ that admits a differentiable extension² $\tilde{f} : \tilde{\mathcal{M}} \rightarrow \mathbb{R}$, where $\tilde{\mathcal{M}}$ is a neighborhood of \mathcal{M} . Then the Riemannian (or projected) gradient simply becomes the orthogonal projection onto $\mathbf{T}_x \mathcal{M}$:

$$\text{grad } f(\mathbf{x}) \stackrel{\text{def}}{=} \Pi_{\mathbf{T}_x \mathcal{M}} \nabla \tilde{f}. \quad (15)$$

In what follows we will often use f instead of \tilde{f} , when it doesn't lead to any confusion.

3.2 Wirtinger calculus and Riemannian gradient

First, we introduce the following real-valued inner product on $\mathbb{C}^{m \times n}$. For two matrices $\mathbf{X}, \mathbf{Y} \in \mathbb{C}^{m \times n}$, we introduce

$$\langle \mathbf{X}, \mathbf{Y} \rangle_{\Re} \stackrel{\text{def}}{=} \langle \mathbf{X}^{\Re}, \mathbf{Y}^{\Re} \rangle + \langle \mathbf{X}^{\Im}, \mathbf{Y}^{\Im} \rangle = \Re(\text{trace}\{\mathbf{X}^H \mathbf{Y}\}), \quad (16)$$

²i.e., f is a restriction of \tilde{f} on \mathcal{M} .

where $\mathbf{X}^{\Re}, \mathbf{Y}^{\Re}$ and $\mathbf{X}^{\Im}, \mathbf{Y}^{\Im}$ are the real and imaginary parts, respectively.

Since a nonconstant function $f : \mathbb{C}^{m \times n} \rightarrow \mathbb{R}$ is never holomorphic, we use a shorthand notation $\frac{\partial f}{\partial \mathbf{X}^{\Re}}, \frac{\partial f}{\partial \mathbf{X}^{\Im}} \in \mathbb{R}^{m \times n}$ for the matrix derivatives with respect to the real and imaginary parts of $\mathbf{X} = \mathbf{X}^{\Re} + i\mathbf{X}^{\Im} \in \mathbb{C}^{m \times n}$. Then the Wirtinger derivatives [1, 10, 24] $\frac{\partial f}{\partial \mathbf{X}}, \frac{\partial f}{\partial \mathbf{X}^*} \in \mathbb{C}^{m \times n}$ are standardly defined as

$$\frac{\partial f}{\partial \mathbf{X}} := \frac{1}{2} \left(\frac{\partial f}{\partial \mathbf{X}^{\Re}} - i \frac{\partial f}{\partial \mathbf{X}^{\Im}} \right), \quad \frac{\partial f}{\partial \mathbf{X}^*} := \frac{1}{2} \left(\frac{\partial f}{\partial \mathbf{X}^{\Re}} + i \frac{\partial f}{\partial \mathbf{X}^{\Im}} \right).$$

The matrix Euclidean gradient of f with respect to (16) becomes

$$\nabla f(\mathbf{X}) = \frac{\partial f}{\partial \mathbf{X}^{\Re}} + i \frac{\partial f}{\partial \mathbf{X}^{\Im}} = 2 \frac{\partial f}{\partial \mathbf{X}^*}(\mathbf{X}).$$

Recall that the unitary group $\mathcal{U}_n \subseteq \mathbb{C}^{n \times n}$ can be considered as an embedded real submanifold of $\mathbb{C}^{n \times n}$ with the inner product induced by (16). By [3, Section 3.5.7], the tangent space to \mathcal{U}_n can be associated with an n^2 -dimensional \mathbb{R} -linear subspace of $\mathbb{C}^{n \times n}$:

$$\mathbf{T}_U \mathcal{U}_n = \{\mathbf{X} \in \mathbb{C}^{n \times n} : \mathbf{X}^H \mathbf{U} + \mathbf{U}^H \mathbf{X} = 0\}.$$

Alternatively, it is the rotated set of skew-Hermitian matrices:

$$\mathbf{T}_U \mathcal{U}_n = \{\mathbf{X} \in \mathbb{C}^{n \times n} : \mathbf{X} = \mathbf{U} \mathbf{Z}, \quad \mathbf{Z} + \mathbf{Z}^H = 0\}.$$

Then for a function $f : \mathbb{C}^{n \times n} \rightarrow \mathbb{R}$, that is differentiable in a neighborhood of $\mathcal{U}_n \subseteq \mathbb{C}^{n \times n}$, by (15), the Riemannian gradient of f becomes.

$$\text{grad } f(\mathbf{U}) = \mathbf{U} \mathbf{\Lambda}(\mathbf{U}), \tag{17}$$

where

$$\mathbf{\Lambda}(\mathbf{U}) = \frac{\mathbf{U}^H \nabla f(\mathbf{U}) - (\nabla f(\mathbf{U}))^H \mathbf{U}}{2} = \mathbf{U}^H \frac{\partial f}{\partial \mathbf{U}^*}(\mathbf{U}) - \left(\frac{\partial f}{\partial \mathbf{U}^*}(\mathbf{U}) \right)^H \mathbf{U}. \tag{18}$$

Note that $\mathbf{\Lambda}(\mathbf{U})$ is a skew-Hermitian matrix, *i.e.*,

$$\Lambda(\mathbf{U})_{i,j} = -(\Lambda(\mathbf{U})_{j,i})^*, \quad 1 \leq i, j \leq n. \tag{19}$$

3.3 Jacobi rotations and directional derivatives

Let f be as in (17) and $h_{(i,j),\mathbf{U}} : \mathcal{U}_2 \rightarrow \mathbb{R}$ be as in (4). Now we compute a “directional derivative” of f corresponding to a Givens transformation.

Lemma 3.1. *The Riemannian gradient of $h_{(i,j),\mathbf{U}}$ at the identity matrix \mathbf{I}_2 is a submatrix of the matrix $\mathbf{\Lambda}(\mathbf{U})$ defined in (18):*

$$\text{grad } h_{(i,j),\mathbf{U}}(\mathbf{I}_2) = \mathcal{P}_{i,j}(\mathbf{\Lambda}(\mathbf{U})) = \begin{bmatrix} \Lambda(\mathbf{U})_{ii} & \Lambda(\mathbf{U})_{ij} \\ \Lambda(\mathbf{U})_{ji} & \Lambda(\mathbf{U})_{jj} \end{bmatrix}. \tag{20}$$

Proof. We denote $h = h_{(i,j),\mathbf{U}}$ for simplicity. Assume $(i, j) = (1, 2)$ such that

$$\mathbf{G}^{(1,2,\Psi)} = \begin{bmatrix} \Psi & \\ & \mathbf{I}_{n-2} \end{bmatrix}$$

for $\Psi \in \mathcal{U}_2$. Define $g(\mathbf{Z}) = f(\mathbf{U}\mathbf{Z})$ for $\mathbf{Z} \in \mathcal{U}_n$. By the composition rule, we have

$$\frac{\partial g}{\partial \mathbf{Z}^*}(\mathbf{Z}) = \mathbf{U}^H \frac{\partial f}{\partial \mathbf{U}^*}(\mathbf{U}\mathbf{Z}).$$

Note that the matrix complex differential [20, Section II] is the first-order approximation, i.e.,

$$h(\mathbf{I}_2 + \Delta\Psi) = h(\mathbf{I}_2) + \left\langle \frac{\partial h}{\partial \Psi}(\mathbf{I}_2), \Delta\Psi \right\rangle_{\Re} + \left\langle \frac{\partial h}{\partial \Psi^*}(\mathbf{I}_2), \Delta\Psi^* \right\rangle_{\Re} + o(\Delta\Psi, \Delta\Psi^*).$$

On the other hand, for $\Delta\mathbf{G}^{(1,2,\Psi)}$, we get that

$$\begin{aligned} h(\mathbf{I}_2 + \Delta\Psi) &= g(\mathbf{I}_n + \begin{bmatrix} \Delta\Psi & \\ & \mathbf{0}_{n-2} \end{bmatrix}) = \\ &= g(\mathbf{I}_n) + \left\langle \frac{\partial g}{\partial \mathbf{Z}}(\mathbf{I}_n), \begin{bmatrix} \Delta\Psi & \\ & \mathbf{0}_{n-2} \end{bmatrix} \right\rangle_{\Re} + \left\langle \frac{\partial g}{\partial \mathbf{Z}^*}(\mathbf{I}_n), \begin{bmatrix} \Delta\Psi^* & \\ & \mathbf{0}_{n-2} \end{bmatrix} \right\rangle_{\Re} + o\left(\Delta\mathbf{G}^{(1,2,\Psi)}, (\Delta\mathbf{G}^{(1,2,\Psi)})^*\right), \end{aligned}$$

from the first-order expansion of g . Hence,

$$\frac{\partial h}{\partial \Psi^*}(\mathbf{I}_2) = \mathcal{P}_{i,j} \left(\frac{\partial g}{\partial \mathbf{Z}^*}(\mathbf{I}_n) \right) = \mathcal{P}_{i,j} \left(\mathbf{U}^H \frac{\partial f}{\partial \mathbf{U}^*}(\mathbf{U}) \right).$$

Then, it follows by (17) that

$$\text{grad } h(\mathbf{I}_2) = \frac{\partial h}{\partial \Psi^*}(\mathbf{I}_2) - \left(\frac{\partial h}{\partial \Psi^*}(\mathbf{I}_2) \right)^H = \mathcal{P}_{i,j}(\Lambda(\mathbf{U})),$$

which completes the proof. \square

3.4 Jacobi-G algorithm

We are now in a position to formulate a general-purpose Jacobi-G algorithm, which is a generalization of the algorithm proposed in [22]. The main ideas behind the algorithm are:

- optimize the cost function by successive Jacobi rotations;
- choose the Jacobi rotations according to a *gradient based order* (well-aligned with the Riemannian gradient).

Algorithm 2. (*General Jacobi-G algorithm*)

Input: A differentiable function $f : \mathcal{U}_n \rightarrow \mathbb{R}^+$ defined in a neighborhood of \mathcal{U}_n , a positive constant $0 < \delta < \sqrt{2}/n$, a starting point \mathbf{U}_0 .

Output: Sequence of iterations \mathbf{U}_k .

- **For** $k = 1, 2, \dots$ until a stopping criterion is satisfied do
- Choose an index pair (i_k, j_k) satisfying

$$\|\text{grad } h_{(i_k, j_k), \mathbf{U}_{k-1}}(\mathbf{I}_2)\| \geq \delta \|\text{grad } f(\mathbf{U}_{k-1})\|. \quad (21)$$

- Find Ψ_k that maximizes $h_k(\Psi) \stackrel{\text{def}}{=} h_{(i_k, j_k), \mathbf{U}_{k-1}}(\Psi)$.

- Update $\mathbf{U}_k = \mathbf{U}_{k-1} \mathbf{G}^{(i_k, j_k, \Psi_k)}$.
- End for

Now we show that it is always possible to choose the index pair such that the inequality (21) is satisfied.

Corollary 3.2. *Let f and $h_{(i,j),\mathbf{U}}$ be as in Lemma 3.1. Then*

$$\max_{1 \leq i < j \leq n} \|\text{grad } h_{(i,j),\mathbf{U}}(\mathbf{I}_2)\| \geq \frac{\sqrt{2}}{n} \|\text{grad } f(\mathbf{U})\|.$$

Proof. By (17) and Lemma 3.1, we see that

$$\begin{aligned} \|\text{grad } f(\mathbf{U})\|^2 &= \|\mathbf{\Lambda}(\mathbf{U})\|^2 = \sum_{i,j=1}^{n,n} |\Lambda(\mathbf{U})_{i,j}|^2 \leq \\ &\leq \sum_{1 \leq i < j \leq n} \|\text{grad } h_{(i,j),\mathbf{U}}(\mathbf{I}_2)\|^2 \leq \frac{n^2}{2} \max_{1 \leq i < j \leq n} \|\text{grad } h_{(i,j),\mathbf{U}}(\mathbf{I}_2)\|^2. \end{aligned}$$

The proof is complete. □

4 Derivatives of the cost functions

In this section, we find the concrete derivatives of the cost functions of interest, which can be used to implement the Jacobi-G algorithm.

4.1 Derivatives for scale-invariant functions

In this section, we consider the functions $f : \mathcal{U}_n \rightarrow \mathbb{R}$ that are invariant with respect to multiplication of columns by numbers on the complex unit circle, i.e.,

$$f(\mathbf{U}) = f(\mathbf{U}\mathbf{S}) \tag{22}$$

for any matrix of the form

$$\mathbf{S} = \begin{bmatrix} e^{i\alpha_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\alpha_n} \end{bmatrix}. \tag{23}$$

For example, the functions in Section 2.2 satisfy (22).

Lemma 4.1. *For a function f satisfying (22) the matrix $\mathbf{\Lambda}(\mathbf{U})$ has zeros on the diagonal for any \mathbf{U} .*

Proof. For the matrix

$$\mathbf{\Omega}_k = k \begin{bmatrix} & & k \\ \mathbf{0} & \vdots & \mathbf{0} \\ \dots & i & \\ \mathbf{0} & & \mathbf{0} \end{bmatrix}. \tag{24}$$

we have that

$$\langle \mathbf{\Omega}_k, \mathbf{\Lambda}(\mathbf{U}) \rangle_{\Re} = \langle \mathbf{U}\mathbf{\Omega}_k, \text{grad } f(\mathbf{U}) \rangle_{\Re} = \left(\frac{d}{dt} f(\mathbf{U}e^{t\mathbf{\Omega}_k}) \right) \Big|_{t=0} = 0.$$

Since $\mathbf{\Lambda}(\mathbf{U})$ is skew-Hermitian, the proof is complete. □

4.2 Derivatives for cost functions expressed via quadratic forms

In this subsection, we find the directional derivatives of the cost functions expressed via quadratic forms (14), as well as the cost functions in Section 2.2.

Lemma 4.2. *Let $h_{(i,j),\mathbf{U}}$ be as in (14) and satisfy (11). Then*

$$\text{grad } h_{(i,j),\mathbf{U}}(\mathbf{I}_2) = 2 \begin{bmatrix} 0 & \Gamma_{12}^{(i,j,\mathbf{U})} + i\Gamma_{13}^{(i,j,\mathbf{U})} \\ -\Gamma_{12}^{(i,j,\mathbf{U})} + i\Gamma_{13}^{(i,j,\mathbf{U})} & 0 \end{bmatrix}.$$

Proof. Denote $h = h_{(i,j),\mathbf{U}}$ and $\Gamma = \Gamma^{(i,j,\mathbf{U})}$ for simplicity. By (19) and Lemma 4.1, we see that $\text{grad } h(\mathbf{I}_2)$ is skew-Hermitian, i.e., it can be decomposed as

$$\text{grad } h(\mathbf{I}_2) = 2\omega_1 \Delta_1 + 2\omega_2 \Delta_2,$$

where $\{\Delta_k\}_{k=1}^4$ is the following orthogonal basis of $\mathbf{T}_{\mathbf{I}_2}\mathcal{U}_2$:

$$\Delta_1 = \begin{bmatrix} 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}, \quad \Delta_2 = \begin{bmatrix} 0 & -\frac{i}{2} \\ -\frac{i}{2} & 0 \end{bmatrix}, \quad \Delta_3 = \begin{bmatrix} i & 0 \\ 0 & 0 \end{bmatrix}, \quad \Delta_4 = \begin{bmatrix} 0 & 0 \\ 0 & i \end{bmatrix}. \quad (25)$$

Since $\|\Delta_1\|^2 = \|\Delta_2\|^2 = 1/2$ and $\|\Delta_3\| = \|\Delta_4\| = 1$, we have

$$\omega_k = \langle \Delta_k, \text{grad } h(\mathbf{I}_2) \rangle_{\Re} = \left(\frac{d}{dt} h(e^{t\Delta_k}) \right) \Big|_{t=0}$$

for $1 \leq k \leq 4$. On the other hand, we have

$$\begin{aligned} h(e^{t\Delta_1}) &= h \left(\begin{bmatrix} \cos \frac{t}{2} & -\sin \frac{t}{2} \\ \sin \frac{t}{2} & \cos \frac{t}{2} \end{bmatrix} \right) = \tilde{h}(\cos t, -\sin t, 0), \\ h(e^{t\Delta_2}) &= h \left(\begin{bmatrix} \cos \frac{t}{2} & -i \sin \frac{t}{2} \\ -i \sin \frac{t}{2} & \cos \frac{t}{2} \end{bmatrix} \right) = \tilde{h}(\cos t, 0, -\sin t), \end{aligned}$$

where $\tilde{h}(\mathbf{v}) = \mathbf{v}^\top \Gamma \mathbf{v}$. Since $\nabla \tilde{h}(\mathbf{v}) = 2\Gamma \mathbf{v}$, we have

$$\omega_1 = -\frac{\partial \tilde{h}}{\partial v_2}(1, 0, 0) = -2\Gamma_{21}, \quad \omega_2 = -\frac{\partial \tilde{h}}{\partial v_3}(1, 0, 0) = -2\Gamma_{31},$$

which completes the proof. \square

Now Lemma 2.3 and Lemma 3.1 allow us to find the Riemannian gradients of all the cost functions in Section 2.2.

Corollary 4.3. *(i) For the cost function in (6), we have that*

$$\Lambda(\mathbf{U}) = 2 \sum_{\ell=1}^L \begin{bmatrix} 0 & W_{12}^{(\ell)}(W_{22}^{(\ell)} - W_{11}^{(\ell)}) & \cdots & W_{1n}^{(\ell)}(W_{nn}^{(\ell)} - W_{11}^{(\ell)}) \\ -W_{21}^{(\ell)}(W_{22}^{(\ell)} - W_{11}^{(\ell)}) & 0 & \cdots & W_{2n}^{(\ell)}(W_{nn}^{(\ell)} - W_{22}^{(\ell)}) \\ \vdots & \vdots & \ddots & \vdots \\ -W_{n1}^{(\ell)}(W_{nn}^{(\ell)} - W_{11}^{(\ell)}) & -W_{n2}^{(\ell)}(W_{nn}^{(\ell)} - W_{22}^{(\ell)}) & \cdots & 0 \end{bmatrix}.$$

(ii) For the cost function in (8), we have that

$$\Lambda(\mathbf{U})_{ij} = 2(\mathcal{W}_{jjj}\mathcal{W}_{jji}^* + \frac{1}{2}\mathcal{W}_{jjj}^*\mathcal{W}_{ijj} - \mathcal{W}_{iii}^*\mathcal{W}_{ijj} - \frac{1}{2}\mathcal{W}_{iii}\mathcal{W}_{jii}^*), \quad 1 \leq i, j \leq n.$$

(iii) For the cost function in (10), we have that

$$\Lambda(\mathbf{U}) = 2 \begin{bmatrix} 0 & \mathcal{V}_{1222} - \mathcal{V}_{2111}^* & \cdots & \mathcal{V}_{1nnn} - \mathcal{V}_{n111}^* \\ \mathcal{V}_{2111} - \mathcal{V}_{1222}^* & 0 & \cdots & \mathcal{V}_{2nnn} - \mathcal{V}_{n222}^* \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{V}_{n111} - \mathcal{V}_{1nnn}^* & \mathcal{V}_{n222} - \mathcal{V}_{2nnn}^* & \cdots & 0 \end{bmatrix}. \quad (26)$$

4.3 Derivatives of other contrast functions

In fact, the form of the Riemannian gradient can be also derived for other cost functions, which can also be called contrast functions [13] [14], and that have the form

$$f(\mathbf{U}) = \sum_{k=1}^n \gamma(\mathbf{u}_k), \quad (27)$$

where $\gamma(\mathbf{u})$ are real-valued.

Remark 4.4. For contrast-like functions, if

$$\gamma(z\mathbf{u}) = \gamma(\mathbf{u}) \text{ for any } \mathbf{u} \in \mathbb{C}^n, z \in \mathbb{C}, |z| = 1, \quad (28)$$

then the function f satisfies the invariance property (22) and the matrix $\Lambda(\mathbf{U})$ has zeros on its diagonal by Lemma 4.1.

Next, we give some hints on how to compute Riemannian gradients for general contrast functions. Note that

$$\frac{\partial f}{\partial \mathbf{U}^*}(\mathbf{U}) = \left(\frac{\partial \gamma}{\partial \mathbf{u}^*}(\mathbf{u}_1), \dots, \frac{\partial \gamma}{\partial \mathbf{u}^*}(\mathbf{u}_n) \right).$$

Therefore, we just need to compute the derivatives for $\gamma(\mathbf{u})$. Now we show how to do that for multilinear forms.

Proposition 4.5. Let $\mathcal{A} \in \mathbb{C}^{n \times \dots \times n}$ be an d th order tensor (not necessarily symmetric), and consider the following multilinear form

$$g_{\mathcal{A}}(\mathbf{u}) = \mathcal{A} \bullet_1 \mathbf{u}^* \cdots \bullet_{d_1} \mathbf{u}^* \bullet_{d_1+1} \mathbf{u} \cdots \bullet_d \mathbf{u}, \quad (29)$$

where $d = d_1 + d_2$ and where $\mathcal{A} \bullet_k \mathbf{v}$ denotes the contraction on the k th index of \mathcal{A} with vector \mathbf{v} . Then the gradients are given by

$$\begin{aligned} \frac{\partial g_{\mathcal{A}}}{\partial \mathbf{u}^*}(\mathbf{u}) &= \sum_{k=1}^{d_1} \mathcal{A} \bullet_1 \mathbf{u}^* \cdots \cancel{\bullet_k \mathbf{u}^*} \cdots \bullet_{d_1} \mathbf{u}^* \bullet_{d_1+1} \mathbf{u} \cdots \bullet_d \mathbf{u}, \\ \frac{\partial g_{\mathcal{A}}}{\partial \mathbf{u}}(\mathbf{u}) &= \sum_{k=1}^{d_2} \mathcal{A} \bullet_1 \mathbf{u}^* \cdots \bullet_{d_1} \mathbf{u}^* \bullet_{d_1+1} \mathbf{u} \cdots \cancel{\bullet_{d_1+k} \mathbf{u}} \cdots \bullet_d \mathbf{u}. \end{aligned}$$

Proof. The result follows from the product differentiation rule and the following identities (see [20, Table IV])

$$\frac{\partial \mathbf{u}^H \mathbf{a}}{\partial \mathbf{u}}(\mathbf{u}) = \mathbf{0}, \quad \frac{\partial \mathbf{u}^H \mathbf{a}}{\partial \mathbf{u}^*} = \mathbf{a}, \quad \frac{\partial \mathbf{u}^T \mathbf{a}}{\partial \mathbf{u}} = \mathbf{a}, \quad \frac{\partial \mathbf{u}^T \mathbf{a}}{\partial \mathbf{u}^*} = \mathbf{0}. \quad \square$$

Proposition 4.6. *For a function*

$$\gamma(\mathbf{u}) = |g_{\mathcal{A}}(\mathbf{u})|^2,$$

with $g_{\mathcal{A}}$ as in (29), the matrix gradient can be computed as:

$$\begin{aligned} \frac{\partial \gamma}{\partial \mathbf{u}^*} &= (g_{\mathcal{A}}(\mathbf{u}))^* \sum_{k=1}^{d_1} \mathcal{A} \bullet_1 \mathbf{u}^* \cdots \cancel{\bullet_k \mathbf{u}^*} \cdots \bullet_{d_1} \mathbf{u}^* \bullet_{d_1+1} \mathbf{u} \cdots \bullet_d \mathbf{u} \\ &+ (g_{\mathcal{A}}(\mathbf{u})) \sum_{k=1}^{d_2} \mathcal{A}^* \bullet_1 \mathbf{u} \cdots \bullet_{d_1} \mathbf{u} \bullet_{d_1+1} \mathbf{u}^* \cdots \cancel{\bullet_{d_1+k} \mathbf{u}^*} \cdots \bullet_d \mathbf{u}^*. \end{aligned}$$

Proof. The result follows³ from Proposition 4.5 and the fact that γ is also a multilinear form:

$$\gamma(\mathbf{u}) = (\mathcal{A} \otimes \mathcal{A}^*) \bullet_1 \mathbf{u}^* \cdots \bullet_{d_1} \mathbf{u}^* \bullet_{d_1+1} \mathbf{u} \cdots \bullet_d \mathbf{u} \bullet_{d+1} \mathbf{u} \cdots \bullet_{d+d_1} \mathbf{u} \bullet_{d+d_1+1} \mathbf{u}^* \cdots \bullet_{2d} \mathbf{u}^*. \quad \square$$

5 Weak convergence results

5.1 Global rates of convergence of descent algorithms on Riemannian manifolds

We first recall the result presented in [8] on convergence of descent algorithms. Although stated initially for retraction-based algorithms, it is valid for any descent algorithms (we provide the sketch of the proof for completeness).

Theorem 5.1 ([8, Theorem 2.5]). *Let $f : \mathcal{M} \rightarrow \mathbb{R}$ be a function bounded from below by f^* . Suppose that, for a sequence of iterations⁴ x_k , there exists $c > 0$ such that the following descent condition holds:*

$$f(x_{k-1}) - f(x_k) \geq c \|\text{grad } f(x_k)\|^2. \quad (30)$$

Then

(i) $\|\text{grad } f(x_k)\| \rightarrow 0$ as $k \rightarrow \infty$;

(ii) We can find an x_k with $\|\text{grad } f(x_k)\| < \varepsilon$ and $f(x_k) \leq f(x_0)$ in at most

$$K_\varepsilon = \left\lceil \frac{f(x_0) - f^*}{c} \frac{1}{\varepsilon^2} \right\rceil$$

iterations; i.e., there exists $k \leq K_\varepsilon$ such that $\|\text{grad } f(x_k)\| < \varepsilon$.

³An alternative proof can be derived by combining Proposition 4.5, the rule of differentiation of composition [20, Theorem 1], and the fact that $d|z|^2 = z^* dz + z dz^*$.

⁴Note that in the original formulation of [8, Theorem 2.5] it was required that x_k are chosen as retractions of some vectors in $\mathbf{T}_{x_{k-1}}$. However, it is easy to see that this condition is not needed in the proof.

Sketch of the proof. The proof follows from the standard inequality for telescopic sums

$$f(x_0) - f^* \geq f(x_0) - f(x_K) = \sum_{k=1}^K f(x_{k-1}) - f(x_k) \geq cK \min_{1 \leq k \leq K} \|\text{grad } f(x_k)\|^2. \quad \square$$

In order to check the descent condition (30), the following lemma about retractions is often useful (we will also use it in this paper).

Definition 5.2. ([3, Definition 4.4.1]) A retraction on a manifold \mathcal{M} is a smooth mapping Retr from the tangent bundle $\mathbf{T}\mathcal{M}$ to \mathcal{M} with the following properties. Let $\text{Retr}_x : \mathbf{T}_x\mathcal{M} \rightarrow \mathcal{M}$ denote the restriction of Retr to the tangent vector space $\mathbf{T}_x\mathcal{M}$.

- (i) $\text{Retr}_x(\mathbf{0}_x) = x$, where $\mathbf{0}_x$ is the zero vector in $\mathbf{T}_x\mathcal{M}$;
- (ii) The differential of Retr_x at $\mathbf{0}_x$, $D\text{Retr}_x(\mathbf{0}_x)$, is the identity map.

Lemma 5.3 ([8, Lemma 2.7]). Let $\mathcal{M} \subseteq \mathbb{R}^n$ be a compact Riemannian submanifold. Let Retr be a retraction on \mathcal{M} . Suppose that $f : \mathcal{M} \rightarrow \mathbb{R}$ has Lipschitz continuous gradient in the convex hull of \mathcal{M} . Then there exists $L \geq 0$ such that for all $x \in \mathcal{M}$ and all $\eta \in \mathbf{T}_x\mathcal{M}$, it holds that

$$|f(\text{Retr}_x(\eta)) - (f(x) + \langle \eta, \text{grad } f(x) \rangle)| \leq \frac{L}{2} \|\eta\|^2. \quad (31)$$

In other words, $f(\text{Retr}_x(\eta))$ is uniformly well approximated by its first order Taylor expansion.

5.2 Convergence of Jacobi-G algorithm to stationary points and global convergence rates

As we will show in this subsection, the iterations in Algorithm 2 can be viewed as a special case of the iterations in Theorem 5.1. Therefore, Theorem 5.1 can be applied and the following proposition on convergence of Jacobi-G algorithm holds true.

Proposition 5.4. Let $f : \mathcal{U}_n \rightarrow \mathbb{R}^+$ have Lipschitz continuous gradient in the convex hull of \mathcal{U}_n . Then there exists $L \geq 0$ such that (31) holds by Lemma 5.3. For Algorithm 2, we have that:

- (i) $\|\text{grad } f(\mathbf{U}_k)\| \rightarrow 0$ in Algorithm 2; in particular, every accumulation point in Algorithm 2 is a stationary point.
- (ii) In order to reach an ε -optimal solution (i.e., $\|\text{grad } f(\mathbf{U}_k)\| \leq \varepsilon$), the algorithm needs at most

$$\left\lceil \frac{2L(f^* - f(x_0))}{\delta^2} \frac{1}{\varepsilon^2} \right\rceil$$

iterations, where δ is from (21).

In order to prove Proposition 5.4, we show that the descent conditions are satisfied using Lemma 5.3.

Lemma 5.5. Let $f : \mathcal{U}_n \rightarrow \mathbb{R}^+$ have Lipschitz continuous gradient in the convex hull of \mathcal{U}_n . Then there exists $L \geq 0$ such that (31) holds by Lemma 5.3. Let $h_{(i,j),\mathbf{U}}$ be as in (4) and Ψ_{opt} be its maximizer. Then

$$h_{(i,j),\mathbf{U}}(\Psi_{\text{opt}}) - h_{(i,j),\mathbf{U}}(\mathbf{I}_2) \geq \frac{\|\text{grad } h_{(i,j),\mathbf{U}}(\mathbf{I}_2)\|^2}{2L}.$$

Proof. Denote $h = h_{(i,j),\mathbf{U}}$ for simplicity. We set

$$\Delta = \mathbf{U} \mathcal{P}_{i,j}^\top \mathcal{P}_{i,j} (\Lambda(\mathbf{U})) \in \mathbf{T}_{\mathbf{U}} \mathcal{U}_n,$$

where $\mathcal{P}_{i,j}^\top : \mathbb{C}^{2 \times 2} \rightarrow \mathbb{C}^{n \times n}$ is the adjoint operator of $\mathcal{P}_{i,j}$. Then

$$\Delta = \mathbf{U} \cdot \begin{array}{c} i \quad j \\ \left[\begin{array}{cccc} \mathbf{0} & \vdots & \vdots & \mathbf{0} \\ \cdots & \Lambda(\mathbf{U})_{ii} & \Lambda(\mathbf{U})_{ij} & \\ \cdots & \Lambda(\mathbf{U})_{ji} & \Lambda(\mathbf{U})_{jj} & \\ \mathbf{0} & & & \mathbf{0} \end{array} \right] \end{array},$$

which is a projection of $\text{grad } f(\mathbf{U})$ onto the tangent space to the submanifold of the matrices of type $\mathbf{U}\mathbf{G}^{(i,j,\Psi)}$. Take the usual exponential map

$$\text{Exp}_{\mathbf{U}}(\mathbf{U}\Omega) = \mathbf{U} \exp(\Omega),$$

which is a retraction (see [3, Proposition 5.4.1]). Denote $\Psi_1 = \text{Exp}_{\mathbf{I}_2}(\frac{1}{L} \text{grad } h(\mathbf{I}_2))$. Then, by Lemma 5.3, we have that

$$\begin{aligned} h(\Psi_1) - h(\mathbf{I}_2) &= f\left(\text{Exp}_{\mathbf{U}}\left(\frac{\Delta}{L}\right)\right) - f(\mathbf{U}) \\ &\geq \left\langle \frac{\Delta}{L}, \text{grad } f(\mathbf{U}) \right\rangle_{\mathbb{R}} - \frac{L}{2} \left\| \frac{\Delta}{L} \right\|^2 = \frac{\|\text{grad } h(\mathbf{I}_2)\|^2}{2L}. \end{aligned}$$

Since $h(\Psi_{\text{opt}}) - h(\mathbf{I}_2) \geq h(\Psi_1) - h(\mathbf{I}_2)$, the proof is completed. \square

Proof of Proposition 5.4. We would like to apply Theorem 5.1 to the function $-f(\mathbf{U})$ (since we are interested in the maximization of $f(\mathbf{U})$). By Lemma 5.5, we have that

$$f(\mathbf{U}_k) - f(\mathbf{U}_{k-1}) = h_k(\Psi_k) - h_k(\mathbf{I}_2) \geq \frac{1}{2L} \|\text{grad } h_k(\mathbf{I}_2)\|^2 \geq \frac{\delta^2}{2L} \|\text{grad } f(\mathbf{U}_{k-1})\|^2,$$

and thus the descent condition (30) holds with the constant $\frac{\delta^2}{2L}$. \square

Corollary 5.6. *Proposition 5.4 applies to the cost functions introduced in Section 2.2.*

6 Convergence of iterations, Łojasiewicz gradient inequality and geodesic convexity

In this section, we gather some known results and preliminaries that will be helpful for the proof of our main results in Section 7.

6.1 Łojasiewicz gradient inequality and speed of convergence

Here we recall the results on convergence of descent algorithms on analytic submanifolds that use Łojasiewicz gradient inequality [26], as presented in [32]. These results were used in [25] to prove the global convergence of Jacobi-G on the orthogonal group.

Definition 6.1 (Łojasiewicz gradient inequality, [31, Definition 2.1]). *Let $\mathcal{M} \subseteq \mathbb{R}^n$ be a Riemannian submanifold of \mathbb{R}^n . The function $f : \mathcal{M} \rightarrow \mathbb{R}$ satisfies a Łojasiewicz gradient inequality at a point $x \in \mathcal{M}$, if there exist $\delta > 0$, $\sigma > 0$ and $\zeta \in (0, \frac{1}{2}]$ such that for all $y \in \mathcal{M}$ with $\|y - x\| < \delta$, it holds that*

$$|f(x) - f(y)|^{1-\zeta} \leq \sigma \|\text{grad } f(x)\|. \quad (32)$$

The following lemma guarantees that (32) is satisfied for the real analytic functions defined on an analytic manifold.

Lemma 6.2 ([31, Proposition 2.2 and Remark 1]). *Let $\mathcal{M} \subseteq \mathbb{R}^n$ be an analytic submanifold⁵ and $f : \mathcal{M} \rightarrow \mathbb{R}$ be a real analytic function. Then for any point $x \in \mathcal{M}$, the function f satisfies a Łojasiewicz gradient inequality (32) for some⁶ $\delta > 0$, $\sigma > 0$ and $\zeta \in (0, \frac{1}{2}]$.*

Łojasiewicz gradient inequality allows for proving convergence of optimization algorithms to a single limit point.

Theorem 6.3 ([31, Theorem 2.3]). *Let $\mathcal{M} \subseteq \mathbb{R}^n$ be an analytic submanifold and $\{x_k : k \in \mathbb{N}\} \subseteq \mathcal{M}$ be a sequence. Suppose that f is real analytic and, for large enough k ,*

(i) *there exists $\sigma > 0$ such that*

$$|f(x_{k+1}) - f(x_k)| \geq \sigma \|\text{grad } f(x_k)\| \|x_{k+1} - x_k\|; \quad (33)$$

(ii) *$\text{grad } f(x_k) = 0$ implies that $x_{k+1} = x_k$.*

Then any accumulation point x_ of $\{x_k : k \in \mathbb{N}\} \subseteq \mathcal{M}$ is the only limit point.*

If, in addition, for some $\kappa > 0$ and for large enough n it holds that

$$\|x_{k+1} - x_k\| \geq \kappa \|\text{grad } f(x_k)\|, \quad (34)$$

then the following convergence rates apply

$$\|x_k - x^*\| \leq C \begin{cases} e^{-ck}, & \text{if } \zeta = \frac{1}{2} \text{ (for some } c > 0), \\ k^{-\frac{\zeta}{1-2\zeta}}, & \text{if } 0 < \zeta < \frac{1}{2}, \end{cases}$$

where ζ is the parameter in (32).

Remark 6.4. *We can relax the conditions of Theorem 6.3 as follows. We can require just that (33) holds for all k such that $\|x_k - x_*\| < \varepsilon$, where x_* is one of the accumulation points of the sequence and ε is some radius of the neighborhood of x_* . This can be verified by inspecting the proof of Theorem 6.3 (see also the proof of [2, Theorem 3.2])*

In the case $\zeta = \frac{1}{2}$, according to Theorem 6.3, the convergence is linear (similarly to the classic results on local convergence of the gradient descent [29, 9]). In the optimization literature, the inequality (32) with $\zeta = \frac{1}{2}$ is often called Polyak-Łojasiewicz⁷ inequality. In the next subsection, we recall some sufficient conditions for Polyak-Łojasiewicz inequality to hold.

⁵See [23, Definition 2.7.1] or [25, Definition 5.1] for a definition of an analytic submanifold.

⁶The values of δ, σ, ζ depend on a specific point.

⁷The inequality (32) with $\zeta = \frac{1}{2}$ goes back to Polyak [29], who used it for proving linear convergence of the gradient descent.

6.2 Riemannian Hessian and stationary points

For a Riemannian manifold \mathcal{M} and a C^2 function $f : \mathcal{M} \rightarrow \mathbb{R}$, the Riemannian Hessian at $x \in \mathcal{M}$ is a bilinear form on $\mathbf{T}_x\mathcal{M}$ that is classically defined as $\text{Hess}_x f[\eta] = \nabla_\eta \text{grad } f(x)$ for $\eta \in \mathbf{T}_x\mathcal{M}$, where ∇ is the Riemannian connection [3, p.105] on \mathcal{M} . Since we are interested in Euclidean submanifolds in this paper, we use the following simpler expression from [4] via the Weingarten operator.

Lemma 6.5 ([4, (8)-(10)]). *For a function \tilde{f} defined in a neighborhood of $\mathcal{M} \subseteq \mathbb{R}^n$. The Riemannian Hessian of f at $x \in \mathcal{M}$ is the map $\mathbf{T}_x\mathcal{M} \rightarrow \mathbf{T}_x\mathcal{M}$ defined as*

$$\text{Hess}_x f[\eta] = \Pi_{\mathbf{T}_x\mathcal{M}} \text{H}_{\tilde{f}}(x)[\eta] + \mathfrak{A}_x(\eta, \Pi_{(\mathbf{T}_x\mathcal{M})^\perp} \nabla \tilde{f}(x)), \quad (35)$$

where $\eta \in \mathbf{T}_x\mathcal{M}$, $\text{H}_{\tilde{f}}(x)$ is the Euclidean Hessian of \tilde{f} at x , $\nabla \tilde{f}$ is the Euclidean gradient, $\Pi_{\mathbf{T}_x\mathcal{M}}$ (resp. $\Pi_{(\mathbf{T}_x\mathcal{M})^\perp}$) denotes the orthogonal projector onto the tangent space $\mathbf{T}_x\mathcal{M}$ (resp. its orthogonal complement), and \mathfrak{A} is the Weingarten operator⁸

$$\begin{aligned} \mathfrak{A}_x : \mathbf{T}_x\mathcal{M} \times (\mathbf{T}_x\mathcal{M})^\perp &\rightarrow \mathbf{T}_x\mathcal{M}, \\ (z, v) &\mapsto \Pi_{\mathbf{T}_x} D_z (\Pi_{\mathbf{T}_x\mathcal{M}}) v, \end{aligned} \quad (36)$$

where D_z is the differential with respect to z .

Note that the Riemannian Hessian (35) contains two terms: one is the projection of the Euclidean Hessian onto the tangent space, and the other is related to the second fundamental form of the manifold, which is computed via the Weingarten operator (some examples for Weingarten operators are given in [4]). We also note that the second term can be also easily interpreted⁹ for the submanifolds given by equality constraints.

Remark 6.6 (see [30, Lemma 5.1]). *Let $\mathcal{M} \subseteq \mathbb{R}^n$ be a k -dimensional submanifold defined locally around x by $n - k$ equations*

$$\mathbf{h}(x) = 0, \text{ where } \mathbf{h}(x) = [h_1(x) \ \cdots \ h_{n-k}(x)]^\top, \ x \in \mathbb{R}^n,$$

such that $\text{rank}\{\mathbf{J}_\mathbf{h}(x)\} = n - k$. Note that the projector on $\mathbf{T}_x\mathcal{M}$ in this case is given by

$$\Pi_{\mathbf{T}_x\mathcal{M}} = \mathbf{I}_n - (\mathbf{J}_\mathbf{h}(x))^\dagger \mathbf{J}_\mathbf{h}(x).$$

Then (35) can be found as (an $n \times n$ matrix)

$$\text{Hess}_x f = \Pi_{\mathbf{T}_x\mathcal{M}} \left(\text{H}_{\tilde{f}}(x) - \sum_{j=1}^{n-k} \mu_j \text{H}_{h_j}(\mathbf{x}) \right) \Pi_{\mathbf{T}_x\mathcal{M}},$$

where the vector $\boldsymbol{\mu} \in \mathbb{R}^{n-k}$ is defined as $\boldsymbol{\mu} = (\mathbf{J}_\mathbf{h}^\top)^\dagger \nabla \tilde{f}(\mathbf{x})$. Note that at a stationary point, $\boldsymbol{\mu}$ is nothing but the vector of Lagrange multipliers.

The Riemannian Hessian gives well-known necessary and sufficient conditions of local extrema¹⁰ (see, for example, [30, Theorem 4.1]).

⁸We take the expression in [4, eqn. (8)] as a definition of Weingarten operator.

⁹This interpretation will not be used in the paper.

¹⁰Usually, these conditions are formulated for local minima and positive definite Hessians, but we formulate them here for local maxima.

- If x is a local maximum of f , then the matrix $\text{Hess}_x f \preceq 0$ (negative semidefinite);
- If $\text{grad } f(x) = 0$ and $\text{Hess}_x f \prec_{\mathbf{T}_x \mathcal{M}} 0$ (negative definite on $\mathbf{T}_x \mathcal{M}$, i.e., $\text{Hess}_x f \preceq 0$ and $\text{rank}\{\text{Hess}_x f\} = \dim(\mathcal{M})$), then f has a strict local maximum at x .

Finally, we distinguish stationary points with nonsingular Riemannian Hessian matrices.

Definition 6.7. *A stationary point ($x \in \mathcal{M}$, $\text{grad } f(x) = 0$) is called non-degenerate if $\text{Hess}_x f$ is nonsingular on $\mathbf{T}_x \mathcal{M}$.*

6.3 Geodesic convexity

We recall the notion of a geodesic convexity [30] on manifolds, which is a generalization of the notion of convexity of sets and functions. In particular, we relate the definiteness of Riemannian Hessians with local convexity/concavity. In this subsection, we assume that $\mathcal{M} \subseteq \mathbb{R}^n$ is a connected Riemannian manifold.

Definition 6.8 ([30, Definition 2.1]). *A set $\mathcal{A} \subseteq \mathcal{M}$ is called geodesically convex, if any two points $x, y \in \mathcal{A}$ are joined by a geodesic lying in \mathcal{A} .*

Definition 6.9 ([30, Definition 2.2]). *Let $\mathcal{A} \subseteq \mathcal{M}$ be geodesically convex. A function $f : \mathcal{A} \rightarrow \mathbb{R}$ is called geodesically convex (resp. concave) on \mathcal{A} , if it is convex¹¹ (resp. concave) when restricted to geodesics.*

We will later on use the following lemma that guarantees the geodesic convexity of sublevel sets.

Lemma 6.10 ([30, Lemma 2.1]). *Let $\mathcal{A} \subseteq \mathcal{M}$ and $f : \mathcal{A} \rightarrow \mathbb{R}$ be geodesically convex. Then for any $x_0 \in \mathcal{A}$, the level set*

$$\{x \mid f(x) \leq f(x_0), x \in \mathcal{A}\}$$

is geodesically convex.

The following two characterizations of geodesic convexity/concavity will be useful for analysis of the convergence properties in Section 7.

Proposition 6.11 ([30, Corollary 4.1]). *For any C^2 function $f : \mathcal{M} \rightarrow \mathbb{R}$, if $\text{grad } f(x) = 0$ and $\text{Hess}_x f \succ_{\mathbf{T}_x \mathcal{M}} 0$ (resp. $\prec_{\mathbf{T}_x \mathcal{M}} 0$), then there exists a neighborhood \mathcal{U} of x , such that f is geodesically convex (resp. concave) on \mathcal{U} .*

6.4 Łojasiewicz inequality at stationary points

It is known, and widely used in optimization (especially in the Euclidean case), that around a strong local maximum the function satisfies the Polyak-Łojasiewicz inequality. In fact, it is also valid for non-degenerate stationary points, as shown in [21]. Here we recall the most general recent result on possibly degenerate stationary points that satisfy the so-called Morse-Bott property (see also [7, p.248]).

¹¹with respect to the arc length parameters.

Definition 6.12 (Local Morse-Bott, [18, Definition 1.5]). Let \mathcal{M} be a C^∞ submanifold and $f : \mathcal{M} \rightarrow \mathbb{R}$ be a C^2 function. Denote the set of stationary points as

$$\text{Crit} f = \{x \in \mathcal{M} : \text{grad} f(x) = 0\}.$$

The function f is said to be Morse-Bott at $x_0 \in \mathcal{M}$ if there exists an open neighborhood $\mathcal{U} \subseteq \mathcal{M}$ of x_0 such that

- (i) $\mathcal{C} = \mathcal{U} \cap \text{Crit} f$ is a relatively open, smooth submanifold of \mathcal{M} ;
- (ii) $\mathbf{T}_{x_0} \mathcal{C} = \text{Ker Hess}_{x_0} f$.

Remark 6.13. (i) If $x_0 \in \mathcal{M}$ is a non-degenerate stationary point, then f is Morse-Bott at x_0 , since $\{x_0\}$ is a zero-dimensional manifold in this case.

(ii) If $x_0 \in \mathcal{M}$ is a degenerate stationary point, then condition (ii) in Definition 6.12 can be rephrased¹² as

$$\text{rank}\{\text{Hess}_{x_0} f\} = \dim \mathcal{M} - \dim \mathcal{C}. \quad (37)$$

For the functions that satisfy the Morse-Bott property, it was recently shown that the Polyak-Lojasiewicz inequality holds true.

Theorem 6.14 ([18, Theorem 3, Corollary 5]). If $\mathcal{U} \subseteq \mathbb{R}^n$ is an open subset and $f : \mathcal{U} \rightarrow \mathbb{R}$ is Morse-Bott at a stationary point x , then there exist $\delta > 0$ and $\sigma > 0$ such that

$$|f(y) - f(x)| \leq \sigma \|\nabla f(y)\|^2,$$

for any $y \in \mathcal{U}$ satisfying $\|y - x\| \leq \delta$.

We can also easily deduce the same result on a smooth manifold \mathcal{M} .

Proposition 6.15. If $\mathcal{U} \subseteq \mathcal{M}$ is an open subset and a C^2 function $f : \mathcal{U} \rightarrow \mathbb{R}$ is Morse-Bott at a stationary point x , then there exist an open neighborhood $\mathcal{V} \subseteq \mathcal{U}$ of x and $\sigma > 0$ such that

$$|f(y) - f(x)| \leq \sigma \|\text{grad} f(y)\|^2,$$

for all $y \in \mathcal{V}$.

Proof. Consider the exponential map $\text{Exp}_x : \mathbf{T}_x \mathcal{M} \rightarrow \mathcal{M}$, which is a local diffeomorphism. Let $\mathcal{W} \subseteq \mathbf{T}_x \mathcal{M}$ be an open subset such that $\text{Exp}_x(\mathcal{W}) = \mathcal{U}$. Let $\hat{f} = f \circ \text{Exp}_x$ be the composite map from \mathcal{W} to \mathbb{R} . Then

$$\nabla \hat{f}(y') = \mathbf{J}_{\text{Exp}_x}^\top(y') \text{grad} f(y), \quad (38)$$

where $y' \in \mathcal{W}$ and $y = \text{Exp}_x(y')$. It follows that Exp_x gives a diffeomorphism between $\text{Crit} f$ and $\text{Crit} \hat{f}$. Since $\text{Hess}_x f = \text{H}_{\hat{f}}(0)$ by [3, Proposition 5.5.5], we have that \hat{f} is Morse-Bott at 0. Therefore, by Theorem 6.14, there exist $\sigma' > 0$, $\sigma > 0$ and an open neighborhood $\mathcal{V} \subseteq \mathcal{U}$ of x such that

$$|f(y) - f(x)| = |\hat{f}(y') - \hat{f}(0)| \leq \sigma' \|\nabla \hat{f}(y')\|^2 \leq \sigma \|\text{grad} f(y)\|^2,$$

for any $y \in \mathcal{V}$, where the last inequality holds because $\mathbf{J}_{\text{Exp}_x}$ is nonsingular in a neighborhood of x . \square

Remark 6.16. For the case of non-degenerate stationary points and C^∞ functions, Proposition 6.15 is proved in [21, Lemma 4.1], which is a simple corollary of Morse Lemma [28, Lemma 2.2]. For C^∞ functions and Morse-Bott functions, Proposition 6.15 (as noted in [18]) is also a simple corollary of Morse-Bott Lemma [5].

¹²due to the fact that $\mathbf{T}_{x_0} \mathcal{C} \subseteq \text{Ker Hess}_{x_0} f$.

7 Convergence results based on Łojasiewicz inequality

Here we consider only the cost functions that are given by

7.1 Weingarten operator and the Riemannian Hessian

We first find the Weingarten map for \mathcal{U}_n (compared with the case of \mathcal{O}_n given in [4]).

Lemma 7.1. *The Weingarten map for \mathcal{U}_n is given by*

$$\mathfrak{A}_U(\mathbf{Z}, \mathbf{V}) = \mathbf{U} \frac{1}{2} (\mathbf{Z}^H \mathbf{V} - \mathbf{V}^H \mathbf{Z}).$$

Proof. We mainly follow the approach of [4]. The projection onto $\mathbf{T}_U \mathcal{U}_n$ is given by

$$\Pi_{\mathbf{T}_U \mathcal{U}_n}(\mathbf{V}) = \frac{1}{2} (\mathbf{V} - \mathbf{U} \mathbf{V}^H \mathbf{U}).$$

By computing the differential of the projector, we have

$$D_{\mathbf{Z}}(\Pi_{\mathbf{T}_U \mathcal{U}_n}) = -\frac{1}{2} (\mathbf{Z} \mathbf{V}^H \mathbf{U} + \mathbf{U} \mathbf{V}^H \mathbf{Z}).$$

Since $\mathbf{Z} \in \mathbf{T}_U \mathcal{U}_n$ and $\mathbf{V} \in (\mathbf{T}_U \mathcal{U}_n)_{\perp}$, we have $\mathbf{Z} \mathbf{U}^H = \mathbf{U} \mathbf{Z}^H$, $\mathbf{U}^H \mathbf{V} = -\mathbf{V}^H \mathbf{U}$, and hence

$$D_{\mathbf{Z}}(\Pi_{\mathbf{T}_U \mathcal{U}_n}) = \frac{1}{2} (\mathbf{Z} \mathbf{U}^H \mathbf{V} - \mathbf{U} \mathbf{V}^H \mathbf{Z}) = \mathbf{U} \frac{1}{2} (\mathbf{Z}^H \mathbf{V} - \mathbf{V}^H \mathbf{Z}).$$

Finally, since $\mathbf{U} \frac{1}{2} (\mathbf{Z}^H \mathbf{V} - \mathbf{V}^H \mathbf{Z}) \in \mathbf{T}_U \mathcal{U}_n$, the proof is complete. \square

Now we show that the Riemannian Hessian for the elementary rotations can be also computed as a submatrix of the Riemannian Hessian.

Proposition 7.2. *Let $h_{(i,j),U}$ be as in (4), the projection operator $\mathcal{P}_{i,j}$ be as in (3), and $\mathcal{P}_{i,j}^{\top}$ be its adjoint operator. Then*

$$\text{Hess}_{\mathbf{I}_2} h_{(i,j),U} = \mathcal{P}_{i,j} \circ ((\mathbf{I}_n \otimes \mathbf{U}^H) \text{Hess}_U f(\mathbf{I}_n \otimes \mathbf{U})) \circ \mathcal{P}_{i,j}^{\top}.$$

Proof. We denote $h = h_{(i,j),U}$ for simplicity. Define $g(\mathbf{Z}) = f(\mathbf{U} \mathbf{Z})$ for $\mathbf{Z} \in \mathcal{U}_n$. It follows from the proof of Lemma 3.1 that

$$\Pi_{\mathbf{T}_{\mathbf{I}_n} \mathcal{U}_n} = (\mathbf{I}_n \otimes \mathbf{U}^H) \Pi_{\mathbf{T}_U \mathcal{U}_n} (\mathbf{I}_n \otimes \mathbf{U}).$$

Therefore, for the first term in (35) we have

$$\begin{aligned} \Pi_{\mathbf{T}_{\mathbf{I}_2} \mathcal{U}_2} \text{H}_h(\mathbf{I}_2) \Pi_{\mathbf{T}_{\mathbf{I}_2} \mathcal{U}_2} &= \mathcal{P}_{i,j} \Pi_{\mathbf{T}_{\mathbf{I}_n} \mathcal{U}_n} \text{H}_g(\mathbf{I}_n) \Pi_{\mathbf{T}_{\mathbf{I}_n} \mathcal{U}_n} \mathcal{P}_{i,j}^{\top} \\ &= \mathcal{P}_{i,j} \circ ((\mathbf{I}_n \otimes \mathbf{U}^H) \Pi_{\mathbf{T}_U \mathcal{U}_n} \text{H}_f(\mathbf{U}) \Pi_{\mathbf{T}_U \mathcal{U}_n} (\mathbf{I}_n \otimes \mathbf{U})) \circ \mathcal{P}_{i,j}^{\top}. \end{aligned}$$

For the second term in (35), we use the expression of Weingarten operator computed in Lemma 7.1. Denote $\mathbf{V}_2 = \Pi_{(T_{\mathbf{I}_2}\mathcal{U}_2)_\perp} \nabla h(\mathbf{I}_2)$ and $\mathbf{V} = \Pi_{(T_{\mathbf{U}}\mathcal{U}_n)_\perp} \nabla f(\mathbf{U})$. Then

$$\begin{aligned} \langle \mathbf{Y}, \mathfrak{A}_{\mathbf{I}_2}(\mathbf{Z}, \mathbf{V}_2) \rangle_{\mathbb{R}} &= \left\langle \mathcal{P}_{i,j}^\top(\mathbf{Y}), \mathcal{P}_{i,j}^\top \left(\frac{1}{2} (\mathbf{Z}^\mathbf{H} \mathbf{V}_2 - \mathbf{V}_2^\mathbf{H} \mathbf{Z}) \right) \right\rangle_{\mathbb{R}} \\ &= \left\langle \mathcal{P}_{i,j}^\top(\mathbf{Y}), \frac{1}{2} \left(\mathcal{P}_{i,j}^\top(\mathbf{Z})^\mathbf{H} \mathcal{P}_{i,j}^\top(\mathbf{V}_2) - \mathcal{P}_{i,j}^\top(\mathbf{V}_2)^\mathbf{H} \mathcal{P}_{i,j}^\top(\mathbf{Z}) \right) \right\rangle_{\mathbb{R}} \\ &= \left\langle \mathcal{P}_{i,j}^\top(\mathbf{Y}), \frac{1}{2} \left(\mathcal{P}_{i,j}^\top(\mathbf{Z})^\mathbf{H} \mathbf{U}^\mathbf{H} \mathbf{V} - \mathbf{V}^\mathbf{H} \mathbf{U} \mathcal{P}_{i,j}^\top(\mathbf{Z}) \right) \right\rangle_{\mathbb{R}} \\ &= \langle \mathbf{U} \mathcal{P}_{i,j}^\top(\mathbf{Y}), \mathfrak{A}_{\mathbf{U}}(\mathcal{P}_{i,j}^\top(\mathbf{Z}), \mathbf{V}) \rangle_{\mathbb{R}}, \end{aligned}$$

which completes the proof. \square

7.2 Riemannian Hessian for the contrast-like functions

Now we show that the Riemannian Hessian of functions satisfying invariance property (22) (with matrix \mathbf{S} given in) is rank-deficient at stationary points.

Lemma 7.3. *Assume that $f : \mathcal{U}_n \rightarrow \mathbb{R}$ satisfies the invariance property (22).*

(i) *If \mathbf{U} is a stationary point of f , then $\text{rank}\{\text{Hess}_{\mathbf{U}}f\} \leq n(n-1)$, and $\text{Hess}_{\mathbf{U}}f[\mathbf{Z}_k] = \mathbf{0}$ for each matrix*

$$\mathbf{Z}_k = [\mathbf{0} \ \cdots \ \mathbf{0} \ i\mathbf{u}_k \ \mathbf{0} \ \cdots \ \mathbf{0}] \in \mathbf{T}_{\mathbf{U}}\mathcal{U}_n, \quad (39)$$

i.e., all the matrices \mathbf{Z}_k are in the kernel of $\text{Hess}_{\mathbf{U}}f$.

(ii) *In general, we have that*

$$\text{Hess}_{\mathbf{U}}f[\mathbf{Z}_k] = \frac{i\mathbf{U}}{2} \begin{bmatrix} 0 & \cdots & 0 & \Lambda_{1,k} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \Lambda_{k-1,k} & 0 & \cdots & 0 \\ -\Lambda_{k,1} & \cdots & -\Lambda_{k,k-1} & 0 & -\Lambda_{k,k+1} & \cdots & \Lambda_{k,n} \\ 0 & \cdots & 0 & \Lambda_{k+1,k} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \Lambda_{n,k} & 0 & \cdots & 0 \end{bmatrix},$$

where $\Lambda = \Lambda(\mathbf{U})$.

Proof. (i) Note that $\mathbf{Z}_k = \mathbf{U}\Omega$, where $\Omega = \Omega_k$ is defined in (24). Recall that the parallel transport on \mathcal{U}_n along the geodesic in the direction $\mathbf{Z}_k = \mathbf{U}\Omega \in \mathbf{T}_{\mathbf{U}}\mathcal{U}_n$ is given [17, (2.18)]¹³, [19, Chap. 2, Ex. A.6] by

$$\tau_{t\mathbf{Z}} : \mathbf{T}_{\mathbf{U}}\mathcal{U}_n \rightarrow \mathbf{T}_{\text{Exp}_{\mathbf{U}}(t\mathbf{Z})}\mathcal{U}_n, \quad \tau_{t\mathbf{Z}}(\mathbf{Y}) = \mathbf{U}e^{\frac{t\Omega}{2}}\mathbf{U}^\mathbf{H}\mathbf{Y}e^{\frac{t\Omega}{2}},$$

and its inverse is then given by

$$\tau_{t\mathbf{Z}}^{-1}(\mathbf{V}) = \mathbf{U}e^{-\frac{t\Omega}{2}}\mathbf{U}^\mathbf{H}\mathbf{V}e^{-\frac{t\Omega}{2}}.$$

¹³Note that [17, (2.18)] was given for \mathcal{O}_n , but it is straightforward to show that the same expression is valid for \mathcal{U}_n , following [19, Chap. 2, Ex. A.6].

From the link between parallel transport [3, (8.1)] and Riemannian connection, we get

$$\begin{aligned}\text{Hess}_{\mathbf{U}}f[\mathbf{Z}_k] &= \nabla_{\mathbf{Z}_k} \text{grad } f(\mathbf{U}) = \left(\frac{d}{dt} \tau_{t\mathbf{Z}}^{-1} (\text{grad } f(\text{Exp}_{\mathbf{U}}(t\mathbf{Z}_k))) \right) \Big|_{t=0} \\ &= \left(\frac{d}{dt} \mathbf{U} e^{-\frac{t\Omega}{2}} \mathbf{U}^{\text{H}} \text{grad } f(\mathbf{U} e^{t\Omega}) e^{-\frac{t\Omega}{2}} \right) \Big|_{t=0}.\end{aligned}$$

Now recall that due to (22), we have $\nabla f(\mathbf{U} e^{t\Omega}) = \nabla f(\mathbf{U}) e^{t\Omega}$, hence

$$\text{grad } f(\mathbf{U} e^{t\Omega}) = \frac{\nabla f(\mathbf{U} e^{t\Omega}) - \mathbf{U} e^{t\Omega} \nabla f(\mathbf{U} e^{t\Omega})^{\text{H}} \mathbf{U} e^{t\Omega}}{2} = \text{grad } f(\mathbf{U}) e^{t\Omega} = 0,$$

when the last equality holds because \mathbf{U} is a stationary point, hence $\text{Hess}_{\mathbf{U}}f[\mathbf{Z}_k] = \mathbf{0}$.

Since $\dim \mathbf{T}_{\mathbf{U}}\mathcal{U}_n = n^2$ and n vectors \mathbf{Z}_k are linearly independent, we get that $\text{rank}\{\text{Hess}_{\mathbf{U}}f\} \leq n(n-1)$.

(ii) In the general case, we get

$$\text{Hess}_{\mathbf{U}}f[\mathbf{Z}_k] = \left(\frac{d}{dt} \mathbf{U} e^{-\frac{t\Omega}{2}} \mathbf{U}^{\text{H}} \text{grad } f(\mathbf{U}) e^{\frac{t\Omega}{2}} \right) \Big|_{t=0} = \frac{\mathbf{U}}{2} (-\Omega \Lambda(\mathbf{U}) + \Lambda(\mathbf{U}) \Omega),$$

which completes the proof. \square

Next, we find the expressions of the Riemannian Hessians in the case of Jacobi rotations. For each pair of indices (i, j) , we define a 2×2 matrix as follows:

$$\mathfrak{D}_{\mathbf{U}}^{(i,j)} = 2 \left(\begin{bmatrix} \Gamma_{2,2} & \Gamma_{2,3} \\ \Gamma_{3,2} & \Gamma_{3,3} \end{bmatrix} - \Gamma_{1,1} \mathbf{I}_2 \right).$$

Lemma 7.4. *Let $h_{(i,j),\mathbf{U}}$ and $\Gamma = \Gamma^{(i,j),\mathbf{U}}$ be as in (14) satisfying (11). Take the basis in $\mathbf{T}_{\mathbf{I}_2}\mathcal{U}_2$ as in (25). Then*

(i) *The 2×2 leading principal submatrix of the Riemannian Hessian of $h_{(i,j),\mathbf{U}}$ is*

$$(\text{Hess}_{\mathbf{I}_2} h_{(i,j),\mathbf{U}})_{1:2,1:2} = \mathfrak{D}_{\mathbf{U}}^{(i,j)}.$$

(ii) *If, in addition $\text{grad } h_{(i,j),\mathbf{U}}(\mathbf{I}_2) = 0$, then the Riemannian Hessian of $h_{(i,j),\mathbf{U}}$ becomes*

$$\text{Hess}_{\mathbf{I}_2} h_{(i,j),\mathbf{U}} = \begin{bmatrix} \mathfrak{D}_{\mathbf{U}}^{(i,j)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Proof. (ii) Follows from (i) and Lemma 7.3, since $\text{Hess}_{\mathbf{I}_2} h_{(i,j),\mathbf{U}}[\Delta_k] = \mathbf{0}$ for $k = 3, 4$.

(i) We denote $h = h_{(i,j),\mathbf{U}}$ for simplicity. Take $\Omega = \alpha_1 \Delta_1 + \alpha_2 \Delta_2$, where $\alpha_1, \alpha_2 \in \mathbb{R}$ satisfy $\alpha_1^2 + \alpha_2^2 = 1$. By [3, (5.32)], we have that

$$\langle \Omega, \text{Hess}_{\mathbf{I}_2} h[\Omega] \rangle_{\mathbb{R}} = \left(\frac{d^2}{dt^2} h(e^{t\Omega}) \right) \Big|_{t=0}.$$

Note that

$$h(e^{t\Omega}) = h\left(\begin{bmatrix} \cos \frac{t}{2} & -(\alpha_1 + i\alpha_2) \sin \frac{t}{2} \\ (\alpha_1 - i\alpha_2) \sin \frac{t}{2} & \cos \frac{t}{2} \end{bmatrix}\right) = \tilde{h}(\cos t, -\alpha_1 \sin t, -\alpha_2 \sin t). \quad (40)$$

It follows that

$$\frac{d}{dt}h(e^{t\Omega}) = -2 \begin{bmatrix} \sin t & \alpha_1 \cos t & \alpha_2 \cos t \end{bmatrix} \Gamma \begin{bmatrix} \cos t & -\alpha_1 \sin t & -\alpha_2 \sin t \end{bmatrix}^\top,$$

and thus

$$\left(\frac{d^2}{dt^2}h(e^{t\Omega})\right)\Big|_{t=0} = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} \mathfrak{D}_{\mathbf{U}}^{(i,j)} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix},$$

which completes the proof. \square

Corollary 7.5. *If \mathbf{I}_2 is a local maximum of $h_{(i,j),\mathbf{U}}$, then the matrix $\mathfrak{D}_{\mathbf{U}}^{(i,j)}$ is negative semidefinite.*

Remark 7.6. *The matrix $\mathfrak{D}_{\mathbf{U}}^{(i,j)}$ is negative definite if and only if*

$$\Gamma_{11} > \lambda_{\max}\left(\begin{bmatrix} \Gamma_{22} & \Gamma_{23} \\ \Gamma_{23} & \Gamma_{33} \end{bmatrix}\right).$$

If, in addition, $\text{grad } h_{(i,j),\mathbf{U}}(\mathbf{I}_2) = 0$, this is equivalent to saying that $\lambda_1(\Gamma) > \lambda_2(\Gamma)$ (i.e., the first two eigenvalues are separated) and $\Gamma_{11} = \lambda_1(\Gamma)$.

7.3 Preliminary lemmas: checking the decrease conditions

In this subsection, we are going to find some sufficient conditions for (33) and (34) to hold in Algorithm 2, which will allow us to use Theorem 6.3.

Let $\mathbf{U}_k = \mathbf{U}_{k-1} \mathbf{G}^{(i_k, j_k, \Psi_k)}$ be the iterations in Algorithm 2. Then it is obvious that

$$\|\mathbf{U}_k - \mathbf{U}_{k-1}\| = \|\Psi_k - \mathbf{I}_2\|.$$

Assume that Ψ_k is obtained as in Section 2.3, i.e., by taking \mathbf{w} as the leading eigenvector of $\Gamma^{(i_k, j_k, \mathbf{U}_{k-1})}$ (normalized so that $w_1 = \cos 2\theta = 2c^2 - 1 > 0$ in (13)) as in Remark 2.4, and retrieving Ψ_k from \mathbf{w} according to (12) and (13). We first express $\|\Psi_k - \mathbf{I}_2\|$ through w_1 .

Lemma 7.7. *For the iterations Ψ_k obtained as in Section 2.3, it holds that*

$$2\|\Psi_k - \mathbf{I}_2\| \geq \sqrt{1 - w_1^2} \geq 0.65\|\Psi_k - \mathbf{I}_2\| \quad (41)$$

Proof. Note that

$$\|\Psi_k - \mathbf{I}_2\| = \sqrt{2(1-c)^2 + 2|s|^2} = 2\sqrt{1-c}.$$

By (13) and Remark 2.4, we see that

$$\sqrt{1 - w_1^2} = 2c\sqrt{1 - c^2} \geq 1.3\sqrt{1 - c}$$

and

$$2c\sqrt{1 - c^2} \leq 2 \cdot 2\sqrt{1 - c}.$$

\square

Since we are looking at Algorithm 2, we can replace $\text{grad } f(\mathbf{U}_{k-1})$ with $\text{grad } h_{(i,j),\mathbf{U}}(\mathbf{I}_2)$ in both inequalities of (41).

Now we prove a result for condition (34).

Lemma 7.8. *Let $f : \mathcal{U}_n \rightarrow \mathbb{R}^+$ be a C^3 function. Then there exists a universal constant $\kappa > 0$ such that*

$$\|\Psi_k - \mathbf{I}_2\| \geq \kappa \|\text{grad } h_k(\mathbf{I}_2)\|.$$

Proof. We denote $\mathbf{\Gamma} = \mathbf{\Gamma}^{(i_k, j_k, \mathbf{U}_{k-1})}$ as in (14). By Lemma 4.2, we have that

$$\|\text{grad } h_k(\mathbf{I}_2)\| = 2\sqrt{2}\sqrt{\Gamma_{12}^2 + \Gamma_{13}^2}.$$

By Lemma 7.7, it is sufficient to prove that

$$1 - w_1^2 \geq \kappa'(\Gamma_{12}^2 + \Gamma_{13}^2)$$

for a universal constant $\kappa' > 0$. Let $\lambda_1 \geq \lambda_2 \geq \lambda_3$ be the eigenvalues of $\mathbf{\Gamma}$. Without loss of generality, we set $\mathbf{\Gamma}' = \mathbf{\Gamma} - \lambda_3 \mathbf{I}_3$, $\mu_1 = \lambda_1 - \lambda_3$ and $\mu_2 = \lambda_2 - \lambda_3$. Then

$$\mathbf{\Gamma}' = \mu_1 \mathbf{w} \mathbf{w}^\top + \mu_2 \mathbf{v} \mathbf{v}^\top, \quad (42)$$

where \mathbf{v} is the second eigenvector of $\mathbf{\Gamma}$. It follows that

$$\begin{aligned} \Gamma_{12}^2 + \Gamma_{13}^2 &= (\Gamma'_{12})^2 + (\Gamma'_{13})^2 = (\mu_1 w_1 w_2 + \mu_2 v_1 v_2)^2 + (\mu_1 w_1 w_3 + \mu_2 v_1 v_3)^2 \\ &= \mu_1^2 w_1^2 (w_2^2 + w_3^2) + 2\mu_1 \mu_2 w_1 v_1 (w_2 v_2 + w_3 v_3) + \mu_2^2 v_1^2 (v_2^2 + v_3^2) \\ &= \mu_1^2 w_1^2 (1 - w_1^2) - 2\mu_1 \mu_2 w_1^2 v_1^2 + \mu_2^2 v_1^2 (1 - v_1^2). \end{aligned} \quad (43)$$

Since $v_1^2 \leq 1 - w_1^2$ (the first elements of the orthogonal eigenvectors) and $\mu_2, \mu_1 \leq 2\|\mathbf{\Gamma}'\|$, we get that

$$\Gamma_{12}^2 + \Gamma_{13}^2 \leq (1 - w_1^2) 4\|\mathbf{\Gamma}'\|^2.$$

By Lemma 4.2, Lemma 7.4 and C^3 smoothness of f , we see that $\mathbf{\Gamma}$ continuously depends on $\mathbf{U} \in \mathcal{U}_n$. Therefore, $\|\mathbf{\Gamma}'\|$ is bounded from above, and thus the proof is completed. \square

We are ready to check the sufficient decrease condition (33).

Lemma 7.9. *Let $\mathbf{\Gamma} = \mathbf{\Gamma}^{(i_k, j_k, \mathbf{U}_{k-1})}$ be as in (14). Let $\lambda_1 \geq \lambda_2 \geq \lambda_3$ be the eigenvalues of $\mathbf{\Gamma}$, and $\eta = \frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3}$. Suppose that there exists $\varepsilon > 0$ such that $1 - \eta \geq \varepsilon$. Then*

$$|h_k(\Psi_k) - h_k(\mathbf{I}_2)| \geq \frac{\varepsilon}{4} \|\text{grad } h_k(\mathbf{I}_2)\| \sqrt{1 - w_1^2}.$$

Proof. Define the ratio

$$q(\mathbf{\Gamma}, \mathbf{w}) = \frac{(\mathbf{w}^\top \mathbf{\Gamma} \mathbf{w} - \Gamma_{11})^2}{(\Gamma_{12}^2 + \Gamma_{13}^2)(1 - w_1^2)}. \quad (44)$$

It is sufficient to prove that $q(\mathbf{\Gamma}, \mathbf{w}) \geq \varepsilon/2$. Now we use the same notations as in the proof of Lemma 7.8. Denote

$$\rho = 1 - w_1^2, \quad \tau = \frac{v_1^2}{\rho} \in [0, 1].$$

From (42), we have that

$$\mathbf{w}^\top \mathbf{\Gamma} \mathbf{w} - \Gamma_{11} = \mu_1 - (\mu_1 w_1^2 + \mu_2 v_1^2) = \mu_1 \rho (1 - \tau \eta). \quad (45)$$

Then, from (43), we have that

$$\Gamma_{12}^2 + \Gamma_{13}^2 = \rho \mu_1^2 ((1 - \tau) + \tau(1 - \eta)^2 - \rho(1 - \tau \eta)^2). \quad (46)$$

By (44), (45) and (46), we see that

$$q(\mathbf{\Gamma}, \mathbf{w}) = \frac{\mu_1^2 \rho^2 (1 - \tau \eta)^2}{\rho^2 \mu_1^2 ((1 - \tau) + \tau(1 - \eta)^2 - \rho(1 - \tau \eta)^2)} \geq \frac{(1 - \tau \eta)^2}{1 - 2\tau \eta + \tau \eta^2}.$$

Note that

$$\frac{1 - 2\tau \eta + \tau \eta^2}{(1 - \tau \eta)^2} = 1 + \frac{(1 - \tau)\tau \eta^2}{(1 - \tau \eta)^2} \leq 1 + \frac{\tau \eta^2}{(1 - \eta)} \leq \frac{2}{\varepsilon}.$$

The proof is complete. \square

7.4 Main results

First, we prove a global convergence result using Theorem 6.3.

Theorem 7.10. *Suppose that $f : \mathcal{U}_n \rightarrow \mathbb{R}^+$ has Lipschitz continuous gradient in the convex hull of \mathcal{U}_n , and $\bar{\mathbf{U}}$ is an accumulation point of Algorithm 2, where the elementary rotations for f are given by (14) (which is a stationary point by Proposition 5.4). Assume that all $\mathfrak{D}_{\mathbf{U}}^{(i,j)}$ are negative definite for all pairs (i, j) at $\bar{\mathbf{U}}$. Then*

(i) $\bar{\mathbf{U}}$ is the only limit point.

(ii) If the rank of Riemannian Hessian is maximal at $\bar{\mathbf{U}}$ (i.e., $\text{rank}\{\text{Hess}_{\bar{\mathbf{U}}} f\} = n^2 - n$), then the speed of convergence is linear.

Moreover, the convergence rates in Theorem 6.3 apply.

Proof. (i) Due to the fact that $\mathfrak{D}_{\mathbf{U}}^{(i,j)}$ is negative definite for any pair (i, j) , the two top eigenvalues of $\mathbf{\Gamma}^{(i,j,\bar{\mathbf{U}})}$ are separated by Remark 7.6. Therefore, there exists $\varepsilon > 0$ such that

$$\frac{\lambda_2(\mathbf{\Gamma}^{(i,j,\bar{\mathbf{U}})}) - \lambda_3(\mathbf{\Gamma}^{(i,j,\bar{\mathbf{U}})})}{\lambda_1(\mathbf{\Gamma}^{(i,j,\bar{\mathbf{U}})}) - \lambda_3(\mathbf{\Gamma}^{(i,j,\bar{\mathbf{U}})})} < 1 - \varepsilon.$$

By the continuity of $\mathbf{\Gamma}^{(i,j,\mathbf{U})}$ with respect to \mathbf{U} , the conditions of Lemma 7.9 are satisfied in a neighborhood of $\bar{\mathbf{U}}$. Therefore, there exists $c > 0$ such that

$$|f(\mathbf{U}_k) - f(\mathbf{U}_{k-1})| \geq c \|\text{grad } h_k(\mathbf{I}_2)\| \|\mathbf{U}_k - \mathbf{U}_{k-1}\|,$$

in a neighborhood of $\bar{\mathbf{U}}$. By Remark 6.4, it is enough to use Theorem 6.3, and thus $\bar{\mathbf{U}}$ is the only limit point. Moreover, by Lemma 7.8, the convergence rates apply.

(ii) Due to the invariance by scaling of the columns, $\bar{\mathbf{U}}$ belongs to an n -dimensional submanifold of stationary points defined by $\bar{\mathbf{U}} \mathbf{S}$, where \mathbf{S} is as in (23). Since $\text{rank}\{\text{Hess}_{\bar{\mathbf{U}}} f\} = n^2 - n$, f is Morse-Bott at $\bar{\mathbf{U}}$ by Remark 6.13. Therefore, by Proposition 6.15, the Łojasiewicz constant $\zeta = 1/2$ in (32) at $\bar{\mathbf{U}}$, and thus the convergence is linear by Theorem 6.3. \square

Theorem 7.11. *Let \mathbf{U}_* be a semi-strict local maximum of f (i.e. such that $\text{rank}\{\text{Hess}_{\mathbf{U}_*} f\} = n^2 - n$). Then there exists a neighborhood \mathcal{W} of \mathbf{U}_* , such that for any $\mathbf{U}_0 \in \mathcal{W}$, Algorithm 2 converges to $\mathbf{U}_* \mathbf{S}$, where \mathbf{S} is of the form (23); the speed of convergence is linear.*

Proof. Let \mathbb{T} be the unit circle in the complex plane and consider the action of the group \mathbb{T}^n on \mathcal{U}_n defined as

$$\mathbf{U} \cdot (t_1, \dots, t_n) = \mathbf{U} \begin{bmatrix} t_1 & & 0 \\ & \ddots & \\ 0 & & t_n \end{bmatrix}.$$

Since the action of \mathbb{T}^n on \mathcal{U}_n is free and proper, we can consider the quotient manifold $\widetilde{\mathcal{U}}_n = \mathcal{U}_n / \mathbb{T}^n$.

In order to define the gradient and Hessians on $\widetilde{\mathcal{U}}_n$, we use the standard splitting into horizontal and vertical space

$$\mathbf{T}_{\mathbf{U}} \mathcal{U}_n = \mathcal{V}_{\mathbf{U}} \mathcal{U}_n \oplus \mathcal{H}_{\mathbf{U}} \mathcal{U}_n,$$

where $\mathcal{H}_{\mathbf{U}} \mathcal{U}_n$ can be found as adding the constraint that the skew-symmetric matrix has zero diagonal:

$$\mathcal{H}_{\mathbf{U}} \mathcal{U}_n = \{\mathbf{X} \in \mathbb{C}^{n \times n} : \mathbf{X} = \mathbf{U} \mathbf{Z}, \quad \mathbf{Z} + \mathbf{Z}^H = 0, \quad \text{diag}\{\mathbf{Z}\} = \mathbf{0}\}.$$

In this case, an element $\widetilde{\mathbf{U}} \in \widetilde{\mathcal{U}}_n$ can be represented as its representative \mathbf{U} and the tangent space $\mathbf{T}_{\widetilde{\mathbf{U}}} \widetilde{\mathcal{U}}_n$ is identified with $\mathcal{H}_{\mathbf{U}} \mathcal{U}_n$, see [3, Section 3.5.8]. Moreover, the Riemannian metric on $\widetilde{\mathcal{U}}_n$ can be defined as

$$\langle \widetilde{\xi}, \widetilde{\eta} \rangle_{\mathbf{T}_{\widetilde{\mathbf{U}}} \widetilde{\mathcal{U}}_n} = \langle \xi, \eta \rangle_{\mathbf{T}_{\mathbf{U}} \mathcal{U}_n},$$

because the inner product is invariant with respect to the choice of representative \mathbf{U} , see [3, Section 3.6.2]. This makes $\widetilde{\mathcal{U}}_n$ a Riemannian manifold; the natural projection $\pi : \mathbf{U} \mapsto \widetilde{\mathbf{U}}$ then becomes a Riemannian submersion.

Due to the invariance property (22), the function f is, in fact, defined on $\widetilde{\mathcal{U}}_n$ (we will denote the corresponding function $\widetilde{f} : \widetilde{\mathcal{U}}_n \rightarrow \mathbb{R}$). Obviously, $\text{grad } f(\mathbf{U}) \in \mathcal{H}_{\mathbf{U}} \mathcal{U}_n$; moreover, the Riemannian Hessian is given by

$$\text{Hess}_{\widetilde{\mathbf{U}}} \widetilde{f}[\mathbf{Z}] = \mathcal{P}_{\mathcal{H}_{\mathbf{U}} \mathcal{U}_n} \text{Hess}_{\mathbf{U}} f[\mathbf{Z}],$$

see [3, Section 5.3.4].

By Lemma 7.3 have that $\text{rank}\{\text{Hess}_{\widetilde{\mathbf{U}}_*} \widetilde{f}\} = \text{rank}\{\text{Hess}_{\mathbf{U}_*} f\} = n(n-1)$, and therefore it is negative definite. Hence, by Proposition 6.11 there exists an open neighborhood $\widetilde{\mathcal{U}}$ of $\widetilde{\mathbf{U}}_*$ where \widetilde{f} is geodesically concave. For simplicity assume that $f(\mathbf{U}_*) = 0$. Without loss of generality (and by Lemma 6.10) we can take $\widetilde{\mathcal{W}}$ such that its boundary $\delta(\widetilde{\mathcal{W}})$ is a level set for a value $a < 0$, i.e.

$$f(\widetilde{\mathbf{U}}) = a, \quad \forall \widetilde{\mathbf{U}} \in \delta(\widetilde{\mathcal{W}}).$$

Moreover, for any $a < b < 0$ we have that the level set

$$\widetilde{\mathcal{W}}_b = \{f(\widetilde{\mathbf{U}}) \geq b, \widetilde{\mathbf{U}} \in \widetilde{\mathcal{W}}\}$$

is a geodesically convex neighbourhood of $\widetilde{\mathbf{U}}_*$.

Next, assume that $\tilde{\mathbf{U}}_{k-1} \in \tilde{\mathcal{W}}$, and consider the $\mathbf{U}_k = \mathbf{U}_{k-1} \mathbf{G}^{(i,j,\Psi)}$ with Ψ given as the maximizer of (14). Define $b = f(\mathbf{U}_{k-1})$. In what follows, we are going to prove that $\tilde{\mathbf{U}}_{k-1} \in \tilde{\mathcal{W}}_b$, so that the sequence $\tilde{\mathbf{U}}_k$ never leaves the set $\tilde{\mathcal{W}}$.

Recall that Ψ is computed as follows (see (2.4)): take the vector \mathbf{w} as in (13). Take $\alpha_1 = -w_2/\sqrt{1-w_1^2}$, $\alpha_2 = -w_3/\sqrt{1-w_1^2}$ (we can assume $w_1 \neq 1$ because otherwise $\Psi = \mathbf{I}_2$ and this case is trivial), and consider the following geodesic in \mathcal{U}_n :

$$\gamma(t) = \mathbf{U}_{k-1} \mathcal{P}_{i,j}^\top \left(\begin{bmatrix} \cos \frac{t}{2} & -(\alpha_1 + i\alpha_2) \sin \frac{t}{2} \\ (\alpha_1 - i\alpha_2) \sin \frac{t}{2} & \cos \frac{t}{2} \end{bmatrix} \right),$$

which starts at $\gamma(0) = \mathbf{U}_{k-1}$. Note that at each point t_1 , $\frac{d}{dt}\gamma(t_1) \in \mathcal{V}_{\gamma(t_1)}\mathcal{U}_n$, hence the corresponding curve $\tilde{\gamma}$ is a geodesic in the quotient manifold $\tilde{\mathcal{U}}_n$.

Next, as in (40), we have that

$$f(\gamma(t)) = [\cos t \quad -\alpha_1 \sin t \quad -\alpha_2 \sin t] \mathbf{\Gamma} [\cos t \quad -\alpha_1 \sin t \quad -\alpha_2 \sin t]^\top + C,$$

hence

$$f(\gamma(t)) = A \cos(2(t - t_*)) + C,$$

where $t_* = \arccos(w_1) \in [0, \frac{\pi}{2}]$ and $\gamma(t_*) = \mathbf{U}_k$. Note that by geodesic concavity of f at \mathbf{U}_{k-1} we have $\frac{d}{dt}f(\gamma(0)) = -4A \cos(-2t_*) < 0$ and therefore $\cos(2t_*) > 0$ and $t_* \in [0, \frac{\pi}{4}]$. Hence we have that $\frac{d}{dt}f(\gamma(t)) = -4A \sin(2(t - t_*)) > 0$ for any $t \in [0, t_*)$, the cost function is decreasing; note that $\frac{d}{dt}f(\gamma(t_*)) = 0$ and there are no other stationary points in $t \in [0, t_*)$.

Next, by continuity and because $\tilde{\mathcal{W}}$ is open, there exist a small $\varepsilon > 0$ such that $\tilde{\gamma}(\varepsilon)$ is in the interior of $\tilde{\mathcal{W}}_b$. By periodicity of $f(\gamma(t))$ and continuity, we have that there exists t_2 such that $\gamma(t_2) \in \delta\tilde{\mathcal{W}}_b$ and $\tilde{\gamma}(t) \in \tilde{\mathcal{W}}_b$ for all $t \in [0, t_2]$. By Rolle's theorem, there exists a local maximum of $f(\gamma(t))$ in $[0, t_2]$. Note that by construction, the closest positive local maximum to 0 is at t_* , therefore $\tilde{\mathbf{U}}_k = \tilde{\gamma}(t_*) \in \tilde{\mathcal{W}}_b$, hence we stay in the same neighbourhood $\tilde{\mathcal{W}}$.

Finally, as a neighborhood of $\mathbf{U}_* \in \mathcal{U}_n$, we can take the preimage $\mathcal{W} = \pi^{-1}(\tilde{\mathcal{W}})$; also linear convergence rate follows from Theorem 7.10. \square

References

- [1] T. E. ABRUDAN, J. ERIKSSON, AND V. KOIVUNEN, *Steepest descent algorithms for optimization under unitary matrix constraint*, IEEE Transactions on Signal Processing, 56 (2008), pp. 1134–1147.
- [2] P. A. ABSIL, R. MAHONY, AND B. ANDREWS, *Convergence of the iterates of descent methods for analytic cost functions*, SIAM Journal on Optimization, 16 (2005), pp. 531–547.
- [3] P.-A. ABSIL, R. MAHONY, AND R. SEPULCHRE, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.
- [4] P. A. ABSIL, R. MAHONY, AND J. TRUMPF, *An extrinsic look at the Riemannian Hessian*, in Geometric Science of Information: First International Conference, GSI 2013, F. Nielsen and F. Barbaresco, eds., Paris, France, 2013, Springer; Berlin Heidelberg, pp. 361–368.

- [5] A. BANYAGA AND D. E. HURTUBISE, *A proof of the Morse-Bott lemma*, *Expositiones Mathematicae*, 22 (2004), pp. 365 – 373.
- [6] W. M. BOOTHBY, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press, 1975.
- [7] R. BOTT, *Nondegenerate critical manifolds*, *Annals of Mathematics*, 60 (1954), pp. 248–261.
- [8] N. BOUMAL, P.-A. ABSIL, AND C. CARTIS, *Global rates of convergence for nonconvex optimization on manifolds*, (2018). preprint: arXiv:1605.08101v2.
- [9] S. BOYD AND L. VANDENBERGHE, *Convex optimization*, Cambridge University Press, 2004.
- [10] D. BRANDWOOD, *A complex gradient operator and its application in adaptive array theory*, *IEE Proceedings H - Microwaves, Optics and Antennas*, 130 (1983), pp. 11–16.
- [11] J.-F. CARDOSO AND A. SOULOUMIAC, *Blind beamforming for non-gaussian signals*, *IEE Proceedings F-Radar and Signal Processing*, 140 (1993), pp. 362–370.
- [12] J.-F. CARDOSO AND A. SOULOUMIAC, *Jacobi angles for simultaneous diagonalization*, *SIAM journal on matrix analysis and applications*, 17 (1996), pp. 161–164.
- [13] P. COMON, *From source separation to blind equalization, contrast-based approaches*, in *Int. Conf. on Image and Signal Processing (ICISP'01)*, Agadir, Morocco, May 2001, pp. 20–32. preprint: hal-01825729.
- [14] P. COMON, *Contrasts, independent component analysis, and blind deconvolution*, *Int. J. Adapt. Control Sig. Proc.*, 18 (2004), pp. 225–243. preprint: hal-00542916.
- [15] P. COMON AND C. JUTTEN, *Handbook of Blind Source Separation: Independent component analysis and applications*, Academic press, 2010.
- [16] L. DE LATHAUWER, *Signal processing based on multilinear algebra*, Katholieke Universiteit Leuven Leuven, 1997.
- [17] A. EDELMAN, T. A. ARIAS, AND S. T. SMITH, *The geometry of algorithms with orthogonality constraints*, *SIAM journal on Matrix Analysis and Applications*, 20 (1998), pp. 303–353.
- [18] P. M. N. FEEHAN, *Optimal Lojasiewicz-Simon inequalities and Morse-Bott Yang-Mills energy functions*, tech. report, 2018. arxiv:1706.09349.
- [19] S. HELGASON, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, 1978.
- [20] A. HJØRUNGNES AND D. GESBERT, *Complex-valued matrix differentiation: Techniques and key results*, *IEEE Transactions on Signal Processing*, 55 (2007), pp. 2740–2746.
- [21] S. HU AND G. LI, *Convergence rate analysis for the higher order power method in best rank one approximations of tensors*, *Numerische Mathematik*, 140 (2018), pp. 993–1031 .

- [22] M. ISHTEVA, P.-A. ABSIL, AND P. VAN DOOREN, *Jacobi algorithm for the best low multilinear rank approximation of symmetric tensors*, SIAM J. Matrix Anal. Appl., 2 (2013), pp. 651–672.
- [23] S. KRANTZ AND H. PARKS, *A Primer of Real Analytic Functions*, Birkhäuser, Boston, 2002.
- [24] S. G. KRANTZ, *Function theory of several complex variables*, vol. 340, American Mathematical Soc., 2001.
- [25] J. LI, K. USEVICH, AND P. COMON, *Globally convergent jacobi-type algorithms for simultaneous orthogonal symmetric tensor diagonalization*, SIAM Journal on Matrix Analysis and Applications, 39 (2018), pp. 1–22.
- [26] S. ŁOJASIEWICZ, *Une propriété topologique des sous ensembles analytiques réels*, in Colloques internationaux du C.N.R.S, 117. Les Équations aux Dérivées Partielles, 1963, pp. 87–89.
- [27] P. MCCULLAGH, *Tensor Methods in Statistics*, Monographs on Statistics and Applied Probability, Chapman and Hall, 1987.
- [28] J. MILNOR, *Morse theory*, Princeton University Press, 1963.
- [29] B. T. POLYAK, *Gradient methods for minimizing functionals*, Zh. Vychisl. Mat. Mat. Fiz., 3 (1963), pp. 643–653.
- [30] T. RAPCSÁK, *Geodesic convexity in nonlinear optimization*, Journal of Optimization Theory and Applications, 69 (1991).
- [31] R. SCHNEIDER AND A. USCHMAJEV, *Convergence results for projected line-search methods on varieties of low-rank matrices via lojasiewicz inequality*, SIAM Journal on Optimization, 25 (2015), pp. 622–646.
- [32] A. USCHMAJEV, *A new convergence proof for the higher-order power method and generalizations*, Pac. J. Optim., 11 (2015), pp. 309–321.