

# PERTURBATION BOUNDS FOR THE MOSTOW AND THE BIPOLAR DECOMPOSITIONS

PRIYANKA GROVER AND PRADIP MISHRA

ABSTRACT. Perturbation bounds for Mostow's decomposition and the bipolar decomposition of matrices have been computed. To do so, expressions for the derivative of the geometric mean of two positive definite matrices have been derived.

## 1. INTRODUCTION

Matrix factorizations have been used in numerical analysis to implement efficient matrix algorithms. In machine learning, matrix factorizations play an important role to explain latent features underlying the interactions between different kinds of entities. Many matrix factorizations namely, the polar decomposition, the  $QR$  decomposition, the  $LR$  decomposition etc., have been of considerable interest for many decades. Perturbation bounds for such factorizations have been of interest for a long time (see [2, 21, 22] and the references therein). Some generalizations and improvements on them have been obtained in the subsequent works, for example, see [11, 12, 13, 16, 17, 18, 19, 23].

An interesting matrix factorization follows from the work of Mostow [20]. It states that every non singular complex matrix  $Z$  can be uniquely factorized as

$$Z = W e^{iK} e^S, \quad (1.1)$$

where  $W$  is a unitary matrix,  $S$  is a real symmetric matrix and  $K$  is a real skew symmetric matrix. Recently, Bhatia [6] showed that every complex unitary matrix  $W$  can be factorized as

$$W = e^L e^{iT}, \quad (1.2)$$

where  $L$  is a real skew symmetric matrix and  $T$  is a real symmetric matrix. Using (1.1) and (1.2), it has been obtained in [6] that

$$Z = e^L e^{iT} e^{iK} e^S. \quad (1.3)$$

Our goal is to find the perturbation bounds for the factors arising in (1.1), (1.2) and (1.3). In [1], Barbaresco has used Berger Fibration in Unit Siegel Disk for Radar

---

2010 *Mathematics Subject Classification.* 15A23, 47A55, 65F60, 47A64, 47A30, 15A45.

*Key words and phrases.* Perturbation bounds, Derivative, Matrix factorizations, The bipolar decomposition, Mostow's decomposition theorem, Polar decomposition, Sylvester's equation, Geometric mean.

Space-Time Adaptive Processing and Toeplitz-Block-Toeplitz covariance matrices based on Mostow's decomposition.

Let  $\mathbb{M}(n, \mathbb{C})$  be the space of  $n \times n$  complex matrices, and  $\mathbb{U}(n, \mathbb{C})$  be the set of  $n \times n$  complex unitary matrices. Let  $||| \cdot |||$  be any unitarily invariant norm on  $\mathbb{M}(n, \mathbb{C})$ , that is, for any  $U, V \in \mathbb{U}(n, \mathbb{C})$  and  $A \in \mathbb{M}(n, \mathbb{C})$ , we have

$$|||UAV||| = |||A|||.$$

Two special examples of such norms are the *operator norm*  $\|\cdot\|$  (also known as the *spectral norm*) and *Frobenius norm*  $\|\cdot\|_2$  (also known as *Hilbert-Schmidt norm* or *Schatten 2-norm*). Various properties of unitarily invariant norms are known [3, Chapter IV]. We would require the following important properties: for  $A, B, C \in \mathbb{M}(n, \mathbb{C})$

$$|||ABC||| \leq \|A\| |||B||| \|C\|, \quad (1.4)$$

and

$$|||A||| = |||A^*||| = |||A^t||| = |||\bar{A}|||. \quad (1.5)$$

Let  $\mathscr{W}$  be a subspace of  $\mathbb{M}(n, \mathbb{C})$  and let  $\mathcal{T} : \mathscr{W} \rightarrow \mathbb{M}(n, \mathbb{C})$  be a linear map. As in [2], we take

$$|||\mathcal{T}||| = \sup\{|||\mathcal{T}(X)||| : |||X||| = 1\}. \quad (1.6)$$

It has been shown in [1, 6] that the factors in the decomposition (1.1) are related to the *geometric mean*. So to obtain the perturbation bounds for (1.1), we obtain expressions for the derivative of the geometric mean and bounds on its norms in Section 2. In Section 3 and Section 4, we exploit the idea in [2] to obtain bounds on the derivative of the decomposition maps for (1.1) and (1.2), respectively. In Section 5, we discuss the first order perturbation bounds for maps on Lie groups and obtain the perturbation bounds for the factorizations (1.1), (1.2) and (1.3).

## 2. DERIVATIVE OF THE GEOMETRIC MEAN

Let  $\mathbb{H}(n, \mathbb{C})$  be the space of  $n \times n$  complex Hermitian matrices and let  $\mathbb{P}(n, \mathbb{C})$  be the set of  $n \times n$  complex positive definite matrices. For  $A, B \in \mathbb{P}(n, \mathbb{C})$  their geometric mean is defined as

$$A\#B = A^{1/2} (A^{-1/2} B A^{-1/2})^{1/2} A^{1/2}, \quad (2.1)$$

[5, Chapter 4]. It is the unique positive solution of the *Riccati equation*

$$XA^{-1}X = B. \quad (2.2)$$

The geometric mean of  $A$  and  $B$  is also given by

$$A\#B = A(A^{-1}B)^{1/2} = (AB^{-1})^{1/2}B, \quad (2.3)$$

where  $(A^{-1}B)^{1/2}$  and  $(AB^{-1})^{1/2}$  are the unique square roots of  $A^{-1}B$  and  $AB^{-1}$ , respectively, with positive eigenvalues.

Let  $G : \mathbb{P}(n, \mathbb{C}) \times \mathbb{P}(n, \mathbb{C}) \rightarrow \mathbb{P}(n, \mathbb{C})$  be the map defined as

$$G(A, B) = A\#B.$$

Since  $A \mapsto A^{1/2}$  is a differentiable function on  $\mathbb{P}(n, \mathbb{C})$ , we get from (2.1) that  $G$  is a differentiable map. The derivative is given by

$$DG(A, B)(X, Y) = \left. \frac{d}{dt} \right|_{t=0} G(A + tX, B + tY) \text{ for all } X, Y \in \mathbb{H}(n, \mathbb{C}).$$

The following proposition gives an expression for  $DG(A, B)$ .

**Proposition 2.1.** *For  $X, Y \in \mathbb{H}(n, \mathbb{C})$*

$$DG(A, B)(X, Y) = \int_0^\infty e^{-t(BA^{-1})^{1/2}} (Y + (BA^{-1})^{1/2} X (A^{-1}B)^{1/2}) e^{-t(A^{-1}B)^{1/2}} dt. \quad (2.4)$$

*Proof.* For sufficiently small  $t$ , by (2.2), we have

$$G(A + tX, B + tY)(A + tX)^{-1}G(A + tX, B + tY) = B + tY. \quad (2.5)$$

Differentiating with respect to  $t$  at 0, we get

$$\begin{aligned} (DG(A, B)(X, Y)) A^{-1}G(A, B) - G(A, B)(A^{-1}XA^{-1})G(A, B) \\ + G(A, B)A^{-1}(DG(A, B)(X, Y)) = Y. \end{aligned}$$

Put  $D = DG(A, B)(X, Y)$  and  $C = A^{-1}G(A, B) = (A^{-1}B)^{1/2}$ . Then the above equation can be rewritten as

$$C^*D + DC = Y + C^*XC. \quad (2.6)$$

This is a well studied Sylvester's equation (see [3, 9]). By [3, Theorem VII.2.3], we obtain

$$DG(A, B)(X, Y) = \int_0^\infty e^{-tC^*} (Y + C^*XC) e^{-tC} dt. \quad (2.7)$$

Substituting  $C = (A^{-1}B)^{1/2}$  in (2.7), we obtain (2.4).  $\square$

Some other expressions for the solution of the Sylvester's equation [3, 10] are known. From these, one can obtain other expressions for  $DG(A, B)(X, Y)$ .

Suppose  $A$  and  $B$  commute. Then  $C = (A^{-1}B)^{1/2}$  is Hermitian. Let  $\lambda_1(C) \geq \dots \geq \lambda_n(C)$  denote the eigenvalues of  $C$ . Using [3, Theorem VII.2.15] for (2.6), we obtain

$$|||DG(A, B)(X, Y)||| \leq \frac{\pi}{4\lambda_n(C)} |||Y + C^*XC|||.$$

By (1.6),  $|||DG(A, B)||| = \sup\{|||DG(A, B)(X, Y)||| : |||(X, Y)||| = 1\}$ , where  $|||(X, Y)||| = \max\{|||X|||, |||Y|||\}$ . So we get

$$|||DG(A, B)||| \leq \frac{\pi}{4\lambda_n(C)} (1 + \|C\|^2). \quad (2.8)$$

By Proposition 2.1, we obtain a better bound for  $|||DG(A, B)|||$ . We mention this in the following corollary for general  $A$  and  $B$ .

**Corollary 2.2.** *For  $A, B \in \mathbb{P}(n, \mathbb{C})$*

$$|||DG(A, B)||| \leq \left( \int_0^\infty \|e^{-t(A^{-1}B)^{1/2}}\|^2 dt \right) (1 + \|(A^{-1}B)^{1/2}\|^2). \quad (2.9)$$

In the case when  $A$  and  $B$  commute,  $\int_0^\infty \|e^{-tC}\|^2 dt = \frac{1}{2\lambda_n(C)}$ . So from (2.9), we obtain

$$|||DG(A, B)||| \leq \frac{1}{2\lambda_n(C)} (1 + \|C\|^2). \quad (2.10)$$

In some other cases, a bound on  $\int_0^\infty \|e^{-tC}\|^2 dt$  is easy to calculate. For example, if  $\lambda_n(\operatorname{Re} C)$  is nonnegative (where  $\operatorname{Re} C = \frac{C+C^*}{2}$ ), then we have

$$\int_0^\infty \|e^{-tC}\|^2 dt \leq \frac{1}{2\lambda_n(\operatorname{Re} C)}.$$

This has been observed in [2, 7].

*Remark 2.3.* We observe that for  $A, B \in \mathbb{P}(n, \mathbb{C})$ ,  $DG(A, B)$  is a positive linear map from  $\mathbb{H}(n, \mathbb{C}) \times \mathbb{H}(n, \mathbb{C})$  to  $\mathbb{M}(n, \mathbb{C})$ . So by [5, Theorem 2.6.3], we obtain

$$\|DG(A, B)\| = \|DG(A, B)(I, I)\|.$$

### 3. MOSTOW'S DECOMPOSITION

The Mostow decomposition theorem (1.1) gives that any non singular matrix  $Z$  can be uniquely factorized as  $Z = We^{iK}e^S$ . Let  $P_1 = e^{iK}$  and  $P_2 = e^S$ . Then  $P_1 \in \mathbb{P}(n, \mathbb{C})$  and  $P_2 \in \mathbb{P}(n, \mathbb{R})$ , where  $\mathbb{P}(n, \mathbb{R})$  stands for the set of  $n \times n$  real positive definite matrices. We also have  $\overline{P_1}P_1 = I$ . Such matrices  $X$  which satisfy  $\overline{X}X = I$  are called *circular* (or *coninvolutary*) [15]. Let  $\mathbb{P}_{circ}$  be the set of circular positive definite matrices. Then  $P_1 \in \mathbb{P}_{circ}$ . Let  $\varrho : \mathbb{GL}(n, \mathbb{C}) \rightarrow \mathbb{U}(n, \mathbb{C}) \times \mathbb{P}_{circ} \times \mathbb{P}(n, \mathbb{R})$  be the map

$$\varrho(Z) = (\varrho_0(Z), \varrho_1(Z), \varrho_2(Z)), \quad (3.1)$$

where  $\varrho_0(Z) = W$ ,  $\varrho_1(Z) = P_1$ , and  $\varrho_2(Z) = P_2$ . Since the factorizations in (1.1) are unique, these maps are well defined. The product map  $\tau(W, P_1, P_2) = WP_1P_2$  is the inverse of  $\varrho$ . For any matrix  $A$ , let  $\operatorname{cond}(A)$  denotes the *condition number* of  $A$ .

**Theorem 3.1.** *For  $Z \in \mathbb{GL}(n, \mathbb{C})$  let  $\beta(Z) = \int_0^\infty \|e^{-t((Z^*Z)^{-1}\overline{Z^*Z})^{1/2}}\|^2 dt$ . Then*

$$|||D\varrho_0(Z)||| \leq \frac{\|P_1^{-1}\| \|P_2^{-1}\|}{2} (1 + \|P_1\|\beta(Z)\operatorname{cond}(Z) (1 + \operatorname{cond}(Z)^4)), \quad (3.2)$$

$$|||D\varrho_1(Z)||| \leq \frac{\operatorname{cond}(P_1) \|P_2^{-1}\|}{2} (1 + \|P_1\|\beta(Z)\operatorname{cond}(Z) (1 + \operatorname{cond}(Z)^4)), \quad (3.3)$$

and

$$|||D\varrho_2(Z)||| \leq \beta(Z)\text{cond}(Z) (1 + \text{cond}(Z)^4). \quad (3.4)$$

*Proof.* We know  $\varrho_2(Z) = e^S = (Z^*Z\#\overline{Z^*Z})^{1/2}$ . Let  $f : (0, \infty) \rightarrow \mathbb{R}$  be defined as  $f(t) = t^{1/2}$ ,  $g : \mathbb{P}(n, \mathbb{C}) \rightarrow \mathbb{P}(n, \mathbb{C}) \times \mathbb{P}(n, \mathbb{C})$  as  $g(A) = (A, \overline{A})$ , and  $h : GL(n, \mathbb{C}) \rightarrow \mathbb{P}(n, \mathbb{C})$  as  $h(Z) = Z^*Z$ . Then  $\varrho_2 = f \circ G \circ g \circ h$ , where  $G$  is the geometric mean map as defined in Section 2. By the chain rule,  $D\varrho_2(Z) = Df(Z^*Z\#\overline{Z^*Z}) \circ DG(Z^*Z, \overline{Z^*Z}) \circ Dg(Z^*Z, \overline{Z^*Z}) \circ Dh(Z)$ .

Now by [3, Theorem X.3.1], we obtain that if  $A \in \mathbb{P}(n, \mathbb{C})$ , then

$$|||Df(A)||| \leq \frac{1}{2} \|A^{-1}\|^{1/2}. \quad (3.5)$$

So

$$|||D\varrho_2(Z)(A)||| \leq \frac{1}{2} \|(Z^*Z\#\overline{Z^*Z})^{-1}\|^{1/2} |||DG(Z^*Z, \overline{Z^*Z})(Z^*A + AZ, \overline{Z^*A + AZ})|||.$$

We know that  $(A\#B)^{-1} = A^{-1}\#B^{-1}$  and  $\|A\#B\| \leq \|A\|^{1/2}\|B\|^{1/2}$ . Therefore

$$|||D\varrho_2(Z)(A)||| \leq \frac{1}{2} \|Z^{-1}\| |||DG(Z^*Z, \overline{Z^*Z})(Z^*A + AZ, \overline{Z^*A + AZ})|||.$$

Let  $C = (Z^*Z)^{-1} (Z^*Z\#\overline{Z^*Z}) = ((Z^*Z)^{-1}\overline{Z^*Z})^{1/2}$ . Then  $\|C\| \leq \text{cond}(Z)^2$ . By (2.9), we obtain

$$|||D\varrho_2(Z)(A)||| \leq \frac{1}{2} \|Z^{-1}\| \beta(Z) (1 + \|C\|^2) \|Z^*A + AZ\|, \quad (3.6)$$

and so

$$|||D\varrho_2(Z)(A)||| \leq \beta(Z) \text{cond}(Z) (1 + \text{cond}(Z)^4) \|A\|. \quad (3.7)$$

Equation (3.4) follows from (3.7).

Let  $\mathbb{SH}(n, \mathbb{R})$  be the space of  $n \times n$  real skew symmetric matrices. The tangent space at any point  $P_1$  is given by  $iP_1^{1/2}\mathbb{SH}(n, \mathbb{R})P_1^{1/2}$ . This follows from [2, p. 258].

Let  $D\varrho(Z) : \mathbb{M}(n, \mathbb{C}) \rightarrow W\mathbb{SH}(n, \mathbb{C}) \oplus iP_1^{1/2}\mathbb{SH}(n, \mathbb{R})P_1^{1/2} \oplus \mathbb{H}(n, \mathbb{R})$  be defined as  $D\varrho(Z)(A) = (WX, iP_1^{1/2}Y_1P_1^{1/2}, Y_2)$ , where  $X \in \mathbb{SH}(n, \mathbb{C})$ ,  $Y_1 \in \mathbb{SH}(n, \mathbb{R})$  and  $Y_2 \in \mathbb{H}(n, \mathbb{R})$ . So we have

$$X^* = -X, \overline{Y_1} = Y_1, Y_1^t = -Y_1, \overline{Y_2} = Y_2, Y_2^t = Y_2. \quad (3.8)$$

The map  $D\varrho(Z)$  is the inverse of  $D\tau(W, P_1, P_2)$ , and so

$$D\varrho_0(Z)(A) = WX, D\varrho_1(Z)(A) = iP_1^{1/2}Y_1P_1^{1/2}, D\varrho_2(Z)(A) = Y_2, \quad (3.9)$$

and

$$D\tau(W, P_1, P_2)(WX, iP_1^{1/2}Y_1P_1^{1/2}, Y_2) = A. \quad (3.10)$$

Also,

$$\begin{aligned}
D\tau(W, P_1, P_2)(WX, iP_1^{1/2}Y_1P_1^{1/2}, Y_2) &= \left. \frac{d}{dt} \right|_{t=0} \tau(We^{tX}, P_1^{1/2}e^{itY_1}P_1^{1/2}, P_2 + tY_2) \\
&= \left. \frac{d}{dt} \right|_{t=0} We^{tX}P_1^{1/2}e^{itY_1}P_1^{1/2}(P_2 + tY_2) \\
&= WX P_1 P_2 + WP_1^{1/2}(iY_1)P_1^{1/2}P_2 + WP_1 Y_2. \quad (3.11)
\end{aligned}$$

By (3.10) and (3.11), we obtain

$$WX P_1 P_2 + WP_1^{1/2}(iY_1)P_1^{1/2}P_2 + WP_1 Y_2 = A, \quad (3.12)$$

that is,

$$X + P_1^{1/2}(iY_1)P_1^{-1/2} = (W^*A - P_1 Y_2)(P_1 P_2)^{-1}. \quad (3.13)$$

Taking conjugate transpose on both the sides and using (3.8), we get

$$-X + P_1^{-1/2}(iY_1)P_1^{1/2} = (P_1 P_2)^{-1}(A^*W - Y_2 P_1). \quad (3.14)$$

Adding (3.13) and (3.14) gives

$$(P_1^{1/2}(iY_1)P_1^{1/2})P_1^{-1} + P_1^{-1}(P_1^{1/2}(iY_1)P_1^{1/2}) = \operatorname{Re}((W^*A - P_1 Y_2)(P_1 P_2)^{-1}).$$

By [3, Theorem VII.2.3], we obtain

$$P_1^{1/2}(iY_1)P_1^{1/2} = \int_0^\infty e^{-tP_1^{-1}} \operatorname{Re}((W^*A - P_1 Y_2)(P_1 P_2)^{-1}) e^{-tP_1^{-1}} dt. \quad (3.15)$$

So

$$\begin{aligned}
\|D\varrho_1(Z)(A)\| &= \|P_1^{1/2}(iY_1)P_1^{1/2}\| \\
&\leq \left( \int_0^\infty \|e^{-tP_1^{-1}}\|^2 dt \right) \| \operatorname{Re}((W^*A - P_1 Y_2)(P_1 P_2)^{-1}) \| \\
&\leq \frac{\|P_1\|}{2} \| (W^*A - P_1 Y_2)(P_1 P_2)^{-1} \| \\
&\leq \frac{\operatorname{cond}(P_1) \|P_2^{-1}\|}{2} (1 + \|P_1\| \beta(Z) \operatorname{cond}(Z) (1 + \operatorname{cond}(Z)^4)) \|A\|.
\end{aligned}$$

The last inequality follows from (3.7) and (3.9). Hence we obtain (3.3).

By (3.12), we also have

$$X P_1 + P_1^{1/2}(iY_1)P_1^{1/2} = (W^*A - P_1 Y_2)P_2^{-1}. \quad (3.16)$$

Again taking conjugate transpose on both the sides and using (3.8), we obtain

$$P_1 X + P_1^{1/2}(iY_1)P_1^{1/2} = P_2^{-1}(A^*W - Y_2 P_1). \quad (3.17)$$

Now, subtracting (3.17) from (3.16), we get

$$X P_1 + P_1 X = 2i \operatorname{Im}((W^*A - P_1 Y_2)P_2^{-1}).$$

Again by [3, Theorem VII.2.3], we get

$$X = \int_0^\infty e^{-tP_1} \operatorname{Im} ((W^*A - P_1Y_2)P_2^{-1}) e^{-tP_1} dt.$$

Therefore

$$\begin{aligned} |||D\varrho_0(Z)(A)||| &= |||X||| \\ &\leq \left( \int_0^\infty \|e^{-tP_1}\|^2 dt \right) |||(W^*A - P_1Y_2)P_2^{-1}||| \\ &\leq \frac{\|P_1^{-1}\| \|P_2^{-1}\|}{2} (1 + \|P_1\| \beta(Z) \operatorname{cond}(Z) (1 + \operatorname{cond}(Z)^4)) |||A|||. \end{aligned}$$

From this, (3.2) follows. □

*Remark 3.2.* We have used in (3.7) that  $\|A\#B\| \leq \|A\|^{1/2}\|B\|^{1/2}$ . Better bounds on  $|||D\varrho_2(Z)(X)|||$  can be found using [8, Theorem 2]. For example, we also have

$$\begin{aligned} |||D\varrho_2(Z)||| &\leq \|Z\| \beta(Z) \|(\overline{Z^*Z})^{-1/4} (Z^*Z)^{-1/2} (\overline{Z^*Z})^{-1/4}\| \\ &\quad (1 + \|Z^{-1}\|^4 \|(\overline{Z^*Z})^{1/4} (Z^*Z)^{1/2} (\overline{Z^*Z})^{1/4}\|^2). \end{aligned}$$

*Remark 3.3.* One can find another bound for  $|||D\varrho_1(Z)|||$  in Theorem 3.1 by using the expression  $e^{iK} = e^{-S}Z^*Ze^{-S}$  given in [6]. This can be expressed as  $\varrho_1(Z) = (\varrho_2(Z)^{-1}(Z^*Z)\varrho_2(Z)^{-1})^{1/2}$ . Using this approach, the factor  $\frac{\operatorname{cond}(P_1)}{2}$  in (3.3) gets replaced by  $\|P_1\|^2$ . By the chain rule, we get

$$D\varrho_1(Z)(A) = Df(P_2^{-1}Z^*ZP_2^{-1})(2\operatorname{Re}(P_2^{-1}Z^*)(AP_2^{-1} - ZP_2^{-1}(D(\varrho_2(Z)(A))P_2^{-1}))),$$

where  $f$  is the square root function. By [3, Theorem X.3.1] and using  $ZP_2^{-1} = WP_1$ , we obtain

$$|||D\varrho_1(Z)||| \leq \|P_1\|^2 \|P_2^{-1}\| (1 + \|P_1\| \beta(Z) \operatorname{cond}(Z) (1 + (\operatorname{cond}(Z))^4)). \quad (3.18)$$

#### 4. DECOMPOSITION OF UNITARY MATRICES

Every complex unitary matrix  $W$  can be factorized as  $W = W_1W_2$ , by the second or third polar decomposition of  $W$ . This decomposition is unique if  $W'W$  doesn't have  $-1$  as an eigenvalue. Let  $\mathbb{U} = \{W \in \mathbb{U}(n, \mathbb{C}) \mid -1 \notin \sigma(W'W)\}$ , where  $\sigma(A)$  denotes the spectrum of  $A$  and  $\mathbb{U}_{\operatorname{sym}^+}$  be the set of  $U \in \mathbb{U}(n, \mathbb{C})$  such that  $U' = U$  and  $U$  has all the eigenvalues in the open right half plane. Let  $\mathbb{O}(n, \mathbb{R})$  be the set of real orthogonal matrices. We define  $\Phi : \mathbb{U} \rightarrow \mathbb{O}(n, \mathbb{R}) \times \mathbb{U}_{\operatorname{sym}^+}$  as  $\Phi(W) = (\Phi_1(W), \Phi_2(W))$ , where  $\Phi_1(W) = W_1$  and  $\Phi_2(W) = W_2$ . The product map  $\Psi : \mathbb{O}(n, \mathbb{R}) \times \mathbb{U}_{\operatorname{sym}^+} \rightarrow \mathbb{U}$  is the inverse of  $\Phi$ .

**Theorem 4.1.** *Let  $\sigma(W_2) = \{e^{i\theta_1}, \dots, e^{i\theta_n}\}$ . Let  $\{a_n\}$  be any  $\ell_1$ -sequence such that for all  $\theta = \theta_i - \theta_j$  ( $1 \leq i, j \leq n$ )*

$$\sum_{n=-\infty}^{\infty} (-1)^n a_n e^{in\theta} = \frac{1}{1 + e^{i\theta}}. \quad (4.1)$$

Then for  $k = 1, 2$

$$\|D\Phi_k(W)\| \leq 2 \left( \sum_{n=-\infty}^{\infty} |a_n| \right). \quad (4.2)$$

*Proof.* The map

$$D\Phi(W) : W \text{SH}(n, \mathbb{C}) \rightarrow W_1 \text{SH}(n, \mathbb{R}) \oplus W_2^{\frac{1}{2}} i\mathbb{H}(n, \mathbb{R}) W_2^{\frac{1}{2}}$$

is an isomorphism and its inverse is  $D\Psi(W_1, W_2)$ . For  $X \in \text{SH}(n, \mathbb{R})$  and  $Y \in i\mathbb{H}(n, \mathbb{R})$

$$\begin{aligned} D\Psi(W_1, W_2)(W_1 X, W_2^{\frac{1}{2}} Y W_2^{\frac{1}{2}}) &= \left. \frac{d}{dt} \right|_{t=0} \Psi \left( W_1 e^{tX}, W_2^{\frac{1}{2}} e^{tY} W_2^{\frac{1}{2}} \right) \\ &= W_1 (X W_2 + W_2^{\frac{1}{2}} Y W_2^{\frac{1}{2}}). \end{aligned}$$

Let  $S \in \text{SH}(n, \mathbb{C})$  be such that  $D\Phi(W)(WS) = (W_1 X, W_2^{\frac{1}{2}} Y W_2^{\frac{1}{2}})$ . Then we have

$$WS = W_1 (X W_2 + W_2^{\frac{1}{2}} Y W_2^{\frac{1}{2}}),$$

that is,

$$W_2 S W_2^{-1} = X + W_2^{1/2} Y W_2^{-1/2}. \quad (4.3)$$

Taking transpose on both the sides of the above equation (4.3) and adding the new equation to (4.3), we get

$$W_2 Y + Y W_2 = W_2^{\frac{3}{2}} S W_2^{-\frac{1}{2}} + W_2^{-\frac{1}{2}} S' W_2^{\frac{3}{2}}. \quad (4.4)$$

By [3, Theorem VII.2.7], we obtain

$$Y = \sum_{n=-\infty}^{\infty} a_n (-1)^n \left( W_2^{-n+\frac{1}{2}} S W_2^{n-\frac{1}{2}} + W_2^{-n-\frac{3}{2}} S' W_2^{n+\frac{3}{2}} \right).$$

This gives

$$\|Y\| \leq 2 \left( \sum_{n=-\infty}^{\infty} |a_n| \right) \|S\|.$$

Therefore

$$\begin{aligned} \|D\Phi_2(W)(WS)\| &= \|Y\| \\ &\leq 2 \left( \sum_{n=-\infty}^{\infty} |a_n| \right) \|WS\|. \end{aligned} \quad (4.5)$$



Equation (4.3) can also be written as

$$W_2^{1/2} S W_2^{-1/2} = W_2^{-1/2} X W_2^{1/2} + Y. \quad (4.6)$$

Taking complex conjugate on both sides of the above equation (4.6) and adding the new equation to (4.6), we get

$$X W_2 + W_2 X = W_2 S - S' W_2. \quad (4.7)$$

By similar calculations as done above, we get

$$|||W_1 X||| \leq 2 \left( \sum_{n=-\infty}^{\infty} |a_n| \right) |||S|||,$$

and so

$$|||D\Phi_1(W)(WS)||| = |||X||| \leq 2 \left( \sum_{n=-\infty}^{\infty} |a_n| \right) |||WS|||. \quad (4.8)$$

Equations (4.5) and (4.8) give the required result.  $\square$

## 5. PERTURBATION BOUNDS

In this section, we discuss first order perturbation bounds for a map from a Lie group to a manifold and use it to obtain perturbation bounds for the decomposition maps. Let  $\mathcal{M} \subseteq GL(n, \mathbb{C})$  be a differentiable manifold. For  $A \in \mathcal{M}$  let  $\tilde{A}$  denote a perturbation of  $A$  in a small neighbourhood of  $A$  in  $\mathcal{M}$ . Suppose  $A = A_1 A_2$ . Then  $\tilde{A}_i$  denote the corresponding factors for  $\tilde{A}$ . Let  $f$  be a smooth function on  $\mathcal{M}$ . If  $\mathcal{M}$  is a convex set, then by Taylor's theorem, we have

$$|||f(\tilde{A}) - f(A)||| \leq |||Df(A)||| |||\tilde{A} - A||| + O(|||\tilde{A} - A|||^2). \quad (5.1)$$

We denote this as

$$|||f(\tilde{A}) - f(A)||| \lesssim |||Df(A)||| |||\tilde{A} - A|||.$$

We also note here that if there is a  $M > 0$  such that  $|||Df(A)||| < M$ , then in a small neighborhood of  $A$ , we have

$$|||f(\tilde{A}) - f(A)||| < M |||\tilde{A} - A|||. \quad (5.2)$$

**5.1. First order perturbation bounds.** The function  $\log$  is well defined for all non singular matrices  $A$  if we choose a branch of logarithm. In this case,  $\exp$  is its inverse. The map  $D \log(A) : \mathbb{M}(n, \mathbb{C}) \rightarrow \mathbb{M}(n, \mathbb{C})$  is given by

$$D \log(A)(X) = \int_0^1 (t(A - I) + I)^{-1} X (t(A - I) + I)^{-1} dt. \quad (5.3)$$

For  $\epsilon > 0$  define  $U_\epsilon = \{X \in \mathbb{M}_n(\mathbb{C}) : \|X\|_2 < \epsilon\}$  and  $V_\epsilon = \exp(U_\epsilon)$ . Let  $G \subseteq GL(n, \mathbb{C})$  be a matrix Lie group with Lie algebra  $\mathcal{G}$  and let  $A_0 \in G$ . Then by

[14, Theorem 2.27], there exists an  $\epsilon > 0$  such that the map  $H : U_\epsilon \cap \mathcal{G} \rightarrow A_0 V_\epsilon \cap G$  defined by  $H(X) = A_0 \exp(X)$  is a bijective map. For  $X \in \mathcal{G}$

$$\begin{aligned} DH(O)(X) &= A_0 D \exp(O)(X) \\ &= A_0 \int_0^1 e^{(1-t)O} X e^{tO} dt \\ &= A_0 X. \end{aligned}$$

This gives

$$\| \| DH(O) \| \| \leq \| A_0 \|. \quad (5.4)$$

Let  $H_1 : A_0 V_\epsilon \rightarrow M(n, \mathbb{C})$  be the map defined as  $H_1(W) = \log(A_0^{-1}W)$ . Note that the restriction of  $H_1$  to  $A_0 V_\epsilon \cap G$  is  $H^{-1}$ , and so  $(H, U_\epsilon \cap \mathcal{G}, A_0 V_\epsilon \cap G)$  is a local chart around  $A_0 \in G$ . For  $A \in A_0 V_\epsilon$

$$\begin{aligned} DH_1(A)(X) &= D \log(A_0^{-1}A)(A_0^{-1}X) \\ &= \int_0^1 (t(A_0^{-1}A - I) + I)^{-1} A_0^{-1}X (t(A_0^{-1}A - I) + I)^{-1} dt \\ &= \int_0^1 (t(A - A_0) + A_0)^{-1} X (t(A - A_0) + A_0)^{-1} A_0 dt. \end{aligned}$$

By Taylor's theorem, we have

$$\begin{aligned} \| \| H_1(\tilde{A}) - H_1(A) \| \| &\lesssim \| \| DH_1(A) \| \| \| \| \tilde{A} - A \| \| \\ &\leq \| A_0 \| \left( \int_0^1 \| (t(A - A_0) + A_0)^{-1} \|^2 dt \right) \| \| \tilde{A} - A \| \|. \end{aligned}$$

In particular, when  $A = A_0$ , we have

$$\| \| H_1(\tilde{A}_0) - H_1(A_0) \| \| \lesssim \| A_0 \| \| A_0^{-1} \|^2 \| \| \tilde{A}_0 - A_0 \| \|. \quad (5.5)$$

Let  $G_1 \subseteq GL(n, \mathbb{C})$  be a differential manifold ( $G_1$  may not be a group) and let  $F : G \rightarrow G_1$  be a smooth map. Then  $F \circ H : U_\epsilon \cap \mathcal{G} \rightarrow G_1$ . Note that  $H_1(A_0) = O$ . Let  $\tilde{S} \in U_\epsilon \cap \mathcal{G}$  be in a small neighbourhood of  $O$ . Then we have

$$\| \| (F \circ H)(\tilde{S}) - (F \circ H)(O) \| \| \lesssim \| \| D(F \circ H)(O) \| \| \| \| \tilde{S} \| \|. \quad (5.6)$$

Therefore

$$\| \| (F \circ H)(\tilde{S}) - (F \circ H)(O) \| \| \lesssim \| \| DF(A_0) \| \| \| \| DH(O) \| \| \| \| \tilde{S} \| \|. \quad (5.7)$$

Let  $\tilde{A}_0 = H(\tilde{S})$ . Using equations (5.4) and (5.7), we get

$$\begin{aligned} \| \| F(\tilde{A}_0) - F(A_0) \| \| &\lesssim \| \| DF(A_0) \| \| \| A_0 \| \| \| \tilde{S} \| \| \\ &= \| \| DF(A_0) \| \| \| A_0 \| \| \| H_1(\tilde{A}_0) - H_1(A_0) \| \|. \end{aligned}$$

By (5.5), we obtain

$$\| \| F(\tilde{A}_0) - F(A_0) \| \| \lesssim \| \| DF(A_0) \| \| \text{cond}(A_0)^2 \| \| \tilde{A}_0 - A_0 \| \|. \quad (5.8)$$

In particular, when  $A_0$  is unitary matrix, we get

$$\| \|F(\tilde{A}_0) - F(A_0)\| \| \lesssim \| \|DF(A_0)\| \| \| \tilde{A}_0 - A_0 \| \| . \quad (5.9)$$

**5.2. Perturbation bounds for the bipolar decomposition.** Equation (5.9) and Theorem 4.1 together give the perturbation bounds for the decomposition (1.2). We state this as a proposition below. The notations are as in Section 4.

**Proposition 5.1.** *For  $W \in \mathbb{U}$  and  $k = 1, 2$*

$$\| \|\tilde{W}_k - W_k\| \| \lesssim 2 \left( \sum_{n=-\infty}^{\infty} |a_n| \right) \| \|\tilde{W} - W\| \| . \quad (5.10)$$

As observed in [6], the expression (1.2) gives both the second and third polar decompositions for a unitary matrix  $W$ . Therefore Theorem 4.6 and Theorem 4.7 in [2] give that for each  $k$

$$\| \|\tilde{W}_k - W_k\| \| \lesssim \left( \int_0^{\infty} \|e^{-tW_2}\|^2 dt \right) \| \|\tilde{W} - W\| \| . \quad (5.11)$$

We see that the bounds obtained in (5.10) are sometimes better than the ones given by (5.11). For example, let  $W = \text{diag}(e^{i\theta}, e^{i\theta})$ , where  $\pi/3 < \theta < \pi/2$ . Then  $W_1 = I$ ,  $W_2 = W$  and  $\int_0^{\infty} \|e^{-tW_2}\|^2 dt = \frac{1}{2 \cos \theta} > 1$ . Let  $a_0 = 1/2$  and  $a_n = 0$  for all  $n \neq 0$ . Then  $\{a_n\} \in \ell_1$  satisfies (4.1) and  $2 \sum_{n=-\infty}^{\infty} |a_n| = 1$ .

If the eigenvalues of  $W_2$  are close to  $i$  or  $-i$ , then the bounds in (5.11) are too large. But the bounds we get in (5.10) depend upon how far the eigenvalues of  $W_2$  lie on the unit circle. We explain this below.

Let  $\Theta = \{\theta_i - \theta_j : e^{i\theta_j} \in \sigma(W_2)\} \subseteq (-\delta, \delta)$ , where  $0 < \delta < \pi$ . We define the function  $f : [-\pi, \pi] \rightarrow \mathbb{C}$  as

$$f(\theta) = \begin{cases} \frac{1}{2} + \frac{i \tan \frac{\delta}{2}}{2(\pi - \delta)}(\theta + \pi) & -\pi \leq \theta \leq -\delta, \\ \frac{1}{2} - i \tan \frac{\theta}{2} & -\delta \leq \theta \leq \delta, \\ \frac{1}{2} + \frac{i}{2} \tan \frac{\delta}{2}(-1 + \frac{\theta - \delta}{(\pi - \delta)}) & \delta \leq \theta \leq \pi. \end{cases}$$

Then  $f$  is periodic and absolutely continuous. Also,

$$\int_{-\pi}^{\pi} |f'(\theta)|^2 d\theta = \frac{\tan^2 \frac{\delta}{2}}{2(\pi - \delta)} + \frac{\tan^2 \frac{\delta}{2}}{12} + \frac{\tan \frac{\delta}{2}}{4}.$$

So  $f' \in L^2[-\pi, \pi]$ . Let the Fourier coefficients of  $f$  be  $b_n$ . For the sequence  $a_n = (-1)^n b_n$  we have  $a_n \in \ell^1$  and for  $\theta \in (-\delta, \delta)$

$$\sum_{n=-\infty}^{\infty} (-1)^n a_n e^{in\theta} = \frac{1}{1 + e^{i\theta}} = \frac{1}{2} - \frac{i}{2} \tan \frac{\theta}{2}. \quad (5.12)$$

From [4, p. 117], we know that

$$\sum_{n=-\infty}^{\infty} |b_n| = \sum_{n=-\infty}^{\infty} |a_n| \leq a_0 + \frac{\pi}{\sqrt{3}} \|f'\|_{L^2}.$$

Therefore

$$2 \sum_{n=-\infty}^{\infty} |a_n| \leq 1 + \frac{\pi}{\sqrt{3}} \sqrt{\frac{2 \tan^2 \frac{\delta}{2}}{(\pi - \delta)} + \frac{\tan^2 \frac{\delta}{2}}{3} + \tan \frac{\delta}{2}}. \quad (5.13)$$

If  $\delta$  is very small, that is, if the eigenvalues of  $W_2$  are very close to each other, then by (5.13), we see that  $2 \sum a_n$  is very close to 1.

We now obtain the perturbation bounds for the factors  $S, K, L, T$  in the bipolar decomposition (1.3). For  $Z \in GL(n, \mathbb{C})$  let

$$k(Z) = \left( \int_0^{\infty} \|e^{-t((Z^*Z)^{-1}\overline{Z^*Z})^{1/2}}\|^2 dt \right) \text{cond}(Z) (1 + \text{cond}(Z)^4),$$

and for  $W$  unitary let

$$C(W) = \int_0^1 \|(t(W - I) + I)^{-1}\|^2 dt.$$

Before stating the theorem, we observe a few things about the decomposition (1.3). For any  $Z \in GL(n, \mathbb{C})$  the matrices  $S$  and  $K$  are unique but  $L$  and  $T$  are not unique. If  $e^L$  and  $e^{iT}$  do not have  $-1$  as an eigenvalue, then we can use principal logarithm to define  $L$  and  $T$  uniquely. But if  $e^L$  or  $e^{iT}$  have  $-1$  as an eigenvalue, then we choose  $\alpha \in [-\pi, 0)$  such that  $e^{i\alpha} \notin \sigma(W_1) \cup \sigma(W_2)$ . A branch of logarithm for which  $\arg z \in [\alpha, \alpha + 2\pi)$  gives unique  $S, K, L$  and  $T$ .

**Theorem 5.2.** *Let  $Z \in GL(n, \mathbb{C})$  be such that  $-1 \notin \sigma(Z^*Z((Z^*Z)^{-1}\#(\overline{Z^*Z})^{-1}))$ . Let  $Z = e^L e^{iT} e^{iK} e^S$ , where  $\sigma(e^{iT}) = \{e^{i\theta_1}, \dots, e^{i\theta_n}\}$ . Let  $\{a_n\}$  be any  $\ell_1$ -sequence such that for all  $\theta = \theta_i - \theta_j$  ( $1 \leq i, j \leq n$ )*

$$\sum_{n=-\infty}^{\infty} (-1)^n a_n e^{in\theta} = \frac{1}{1 + e^{i\theta}}.$$

Then

$$\| \|\tilde{L} - L\| \| \lesssim 2 C(e^L) \left( \sum_{n=-\infty}^{\infty} |a_n| \right) \frac{\|e^{-iK}\| \|e^{-S}\|}{2} (1 + \|e^{iK}\| k(Z)) \| \|\tilde{Z} - Z\| \|, \quad (5.14)$$

$$\| \|\tilde{T} - T\| \| \lesssim 2 C(e^{iT}) \left( \sum_{n=-\infty}^{\infty} |a_n| \right) \frac{\|e^{-iK}\| \|e^{-S}\|}{2} (1 + \|e^{iK}\| k(Z)) \| \|\tilde{Z} - Z\| \|, \quad (5.15)$$

$$\| \|\tilde{K} - K\| \| \lesssim \| \|e^{-iK}\| \| \frac{\text{cond}(e^{iK}) \|e^{-S}\|}{2} (1 + \|e^{iK}\| k(Z)) \| \|\tilde{Z} - Z\| \|, \quad (5.16)$$

and

$$|||\tilde{S} - S||| \lesssim |||e^{-S}||| k(Z) |||\tilde{Z} - Z|||. \quad (5.17)$$

*Proof.* Let  $Z = We^{iK}e^S$ . Using notations as in section 3, we get

$$|||\tilde{W} - W||| \lesssim |||\mathbf{D}\varrho_0(Z)||| |||\tilde{Z} - Z|||, \quad (5.18)$$

$$|||e^{i\tilde{K}} - e^{iK}||| \lesssim |||\mathbf{D}\varrho_1(Z)||| |||\tilde{Z} - Z|||, \quad (5.19)$$

and

$$|||e^{\tilde{S}} - e^S||| \lesssim |||\mathbf{D}\varrho_2(Z)||| |||\tilde{Z} - Z|||. \quad (5.20)$$

The matrices  $e^S$  and  $e^{iK}$  are both positive. So  $\log(e^S) = S$  and  $\log(e^{iK}) = iK$ . We have

$$\begin{aligned} |||\tilde{S} - S||| &\lesssim |||\mathbf{D}\log(e^S)||| |||e^{\tilde{S}} - e^S||| \\ &\lesssim |||\mathbf{D}\log(e^S)||| |||\mathbf{D}\varrho_2(Z)||| |||\tilde{Z} - Z|||. \end{aligned} \quad (5.21)$$

Similarly,

$$|||\tilde{K} - K||| \lesssim |||\mathbf{D}\log(e^{iK})||| |||\mathbf{D}\varrho_1(Z)||| |||\tilde{Z} - Z|||. \quad (5.22)$$

We know that if  $\eta$  is an operator monotone function on  $(0, \infty)$ , then for  $A \in \mathbb{P}(n, \mathbb{C})$   $|||\mathbf{D}\eta(A)||| \leq \|\eta'(A)\|$  [3, Theorem X.3.4]. Now since  $\log$  is an operator monotone function on  $(0, \infty)$ , we obtain

$$|||\mathbf{D}\log(e^S)||| \leq |||e^{-S}||| \quad (5.23)$$

and

$$|||\mathbf{D}\log(e^{iK})||| \leq |||e^{-iK}|||. \quad (5.24)$$

Equations (5.23) and (3.4) give (5.17). And, (5.24) and (3.3) give (5.16).

Since  $-1 \notin \sigma(Z'Z((Z^*Z)^{-1}\#(\overline{Z^*Z})^{-1}))$ ,  $-1 \notin \sigma(W'W)$ . Then by the second or third polar decomposition,  $W$  can be uniquely factorized as  $W = W_1W_2$ , where  $W_1$  and  $W_2$  are also unitary matrices. If  $L = \log W_1$  and  $iT = \log W_2$ , then we have  $W = e^L e^{iT}$ . Now

$$\begin{aligned} |||\tilde{L} - L||| &= |||\log(e^{\tilde{L}}) - \log(e^L)||| \\ &\lesssim |||\mathbf{D}\log(e^L)||| |||e^{\tilde{L}} - e^L|||. \end{aligned}$$

By (5.3), we obtain

$$|||\tilde{L} - L||| \lesssim C(e^L) |||e^{\tilde{L}} - e^L|||.$$

Equation (5.10) gives

$$|||\tilde{L} - L||| \lesssim 2C(e^L) \left( \sum_{n=-\infty}^{\infty} |a_n| \right) |||\tilde{W} - W|||.$$

Therefore we have

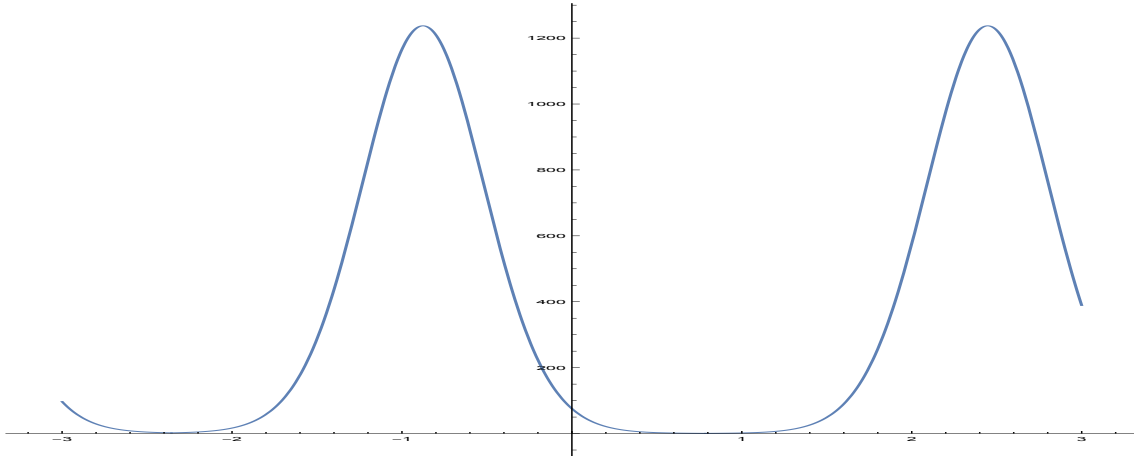
$$|||\tilde{L} - L||| \lesssim 2C(e^L) \left( \sum_{n=-\infty}^{\infty} |a_n| \right) |||\mathbf{D}\varrho_0(Z)||| |||\tilde{Z} - Z|||.$$

By (3.2), we obtain (5.14). Similar calculations by putting  $\tilde{T} = \frac{1}{i} \log \tilde{W}_2$  yield (5.15).  $\square$

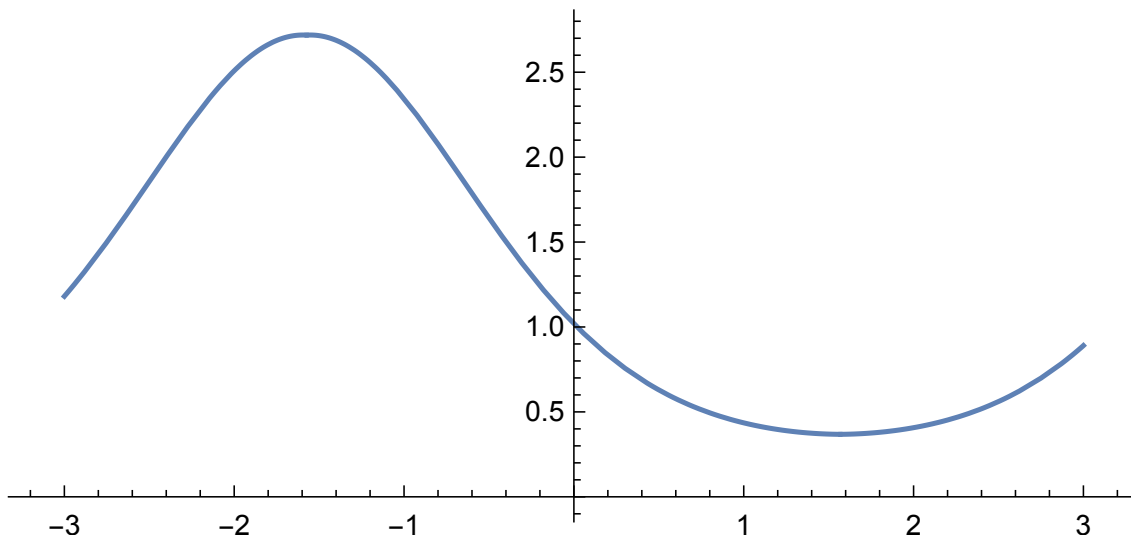
We illustrate the behaviour of the above bounds with the help of an example. For a natural number  $n$ , consider a one-parameter family of matrices  $Z_n(t) = \text{diag}(e^{\sin t}, e^{\sin(t+\frac{\pi}{n})})$ ,  $t \in \mathbb{R}$ . For  $Z_n(t)$ , the factors in the bipolar decomposition are given by  $S_n(t) = \text{diag}(\sin t, \sin(t+\frac{\pi}{n}))$ ,  $K_n(t) = T_n(t) = L_n(t) = O$ . We consider the operator norm in Theorem 5.2. Let  $f_n(t)$  be the first order perturbation bounds as given in (5.17), that is,  $f_n(t) = \|e^{-S_n(t)}\| k(Z_n(t))$ . Then

$$f_n(t) = \frac{1}{2} \left( \max(e^{-\sin t}, e^{-\sin(t+\frac{\pi}{n})}) \right)^2 \max(e^{\sin t}, e^{\sin(t+\frac{\pi}{n})}) \\ (1 + \max(e^{-4\sin t}, e^{-4\sin(t+\frac{\pi}{n})}) \max(e^{4\sin t}, e^{4\sin(t+\frac{\pi}{n})})).$$

The behavior of  $f_n(t)$  can be seen in the following graph. We observe that for  $n = 2$ , the perturbation bound for some of these matrices can be more than 1200.



When  $n$  increases, the maximum value of  $f_n(t)$  decreases. In particular, we observe this for  $n = 500$  in the below graph.



Bounds for other factors  $K_n(t)$ ,  $L_n(t)$  and  $T_n(t)$  are given by  $g_n(t) := \frac{1}{2} \|e^{-S_n(t)}\| (1 + k(Z_n(t)))$  which also vary in a similar way.

*Remark 5.3.* Other perturbation bounds for  $L$  and  $T$  in Theorem 5.2 can also be found using direct formulas, which we get from the principal logarithm. Let  $\mathbb{V}$  be the set of complex unitary matrices  $W$  such that  $W'W$  and  $We^{-\frac{1}{2}\log(W'W)}$  do not have eigenvalue  $-1$ . Then  $L = \log(We^{-\frac{1}{2}\log(W'W)})$  and  $T = \frac{1}{2i} \log(W'W)$ . Using the chain rule and Taylor's theorem, we get

$$|||\tilde{L} - L||| \leq (1 + C(W'W)) |||\tilde{W} - W|||$$

and

$$|||\tilde{T} - T||| \leq C(W'W) |||\tilde{W} - W|||.$$

**Acknowledgement.** The work of the first author is supported by the research grant of INSPIRE Faculty Award [DST/INSPIRE/04/2014/002705] of Department of Science and Technology, India.

#### REFERENCES

- [1] F. Barbaresco, Information geometry of covariance matrix: Cartan-Siegel homogeneous bounded domains, Mostow/Berger fibration and Fréchet median, *Matrix Information Geometry*, eds. F. Nielsen, R. Bhatia, Springer, 2013, pp. 199-255.
- [2] R. Bhatia, Matrix factorizations and their perturbations, *Linear Algebra Appl.* 197/198 (1994) 245-276.
- [3] R. Bhatia, *Matrix Analysis*, Springer, 1997.
- [4] R. Bhatia, *Fourier Series*, Hindustan Book Agency, 2003.
- [5] R. Bhatia, *Positive Definite Matrices*, Princeton University Press, 2007.
- [6] R. Bhatia, The bipolar decomposition, *Linear Algebra Appl.* 439 (2013) 3031-3037.
- [7] R. Bhatia, L. Elsner, Higher order logarithmic derivatives of matrices in the spectral norm, *SIAM J. Matrix Anal. Appl.* 25 (2003) 662-668.

- [8] R. Bhatia, P. Grover, Norm inequalities related to the matrix geometric mean, *Linear Algebra Appl.* 437 (2012) 726–733.
- [9] R. Bhatia, P. Rosenthal, How and why to solve the operator equation  $AX - XB = Y$ , *Bull. London Math. Soc.* 29 (1997) 1–21.
- [10] R. Bhatia, M. Uchiyama, The operator equation  $\sum_{i=0}^n A^{n-i}XB^i = Y$ , *Expo. Math.* 27 (2009) 251–255.
- [11] X.-W. Chang, On the sensitivity of the SR decomposition, *Linear Algebra Appl.* 282 (1998) 297–310.
- [12] X.-W. Chang, D. Stehle, Rigorous perturbation bounds of some matrix factorizations, *SIAM J. Matrix Anal. Appl.* 31 (2010) 2841–2859.
- [13] A. Galántai, Perturbations of triangular matrix factorizations, *Linear Multilinear Algebra* 51 (2003) 175–198.
- [14] B.C. Hall, *Lie Groups, Lie Algebras, and Representations : An Elementary Introduction*, Springer, 2003.
- [15] R. Horn, C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1990.
- [16] A. Largillier, Bounds for relative errors of complex matrix factorizations, *Appl. Math. Lett.* 9 (1996) 79–84.
- [17] W. Li, W. Sun, Perturbation bounds of unitary and subunitary polar factors, *SIAM J. Matrix Anal. Appl.* 23 (2002) 1183–1193.
- [18] W. Li, W. Sun, New perturbation bounds for unitary polar factors, *SIAM J. Matrix Anal. Appl.* 25 (2003) 362–372.
- [19] H. Li, H. Yang, H. Shao, Perturbation analysis for the hyperbolic QR factorization, *Comput. Math. Appl.* 63 (2012) 1607–1620.
- [20] G.D. Mostow, Some new decomposition theorems for semi-simple groups, *Mem. Amer. Math. Soc.* 14 (1955) 31–54.
- [21] G.W. Stewart, Perturbation bounds for the QR factorization of a matrix, *SIAM J. Numer. Anal.* 14 (1977) 509–518.
- [22] G.W. Stewart, J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, 1990.
- [23] Z. Xie, W. Li, Sensitivity analysis for the SR decomposition, *Linear Multilinear Algebra* 63 (2015) 222–234.

*E-mail address:* priyanka.grover@snu.edu.in, pradip.kumar@snu.edu.in

DEPARTMENT OF MATHEMATICS, SHIV NADAR UNIVERSITY, DADRI, U.P. 201314, INDIA.