Ann. Funct. Anal. 9 (2018), no. 4, 435–450
https://doi.org/10.1215/20088752-2017-0059
ISSN: 2008-8752 (electronic)
http://projecteuclid.org/afa

PERTURBATION ANALYSIS FOR THE (SKEW) HERMITIAN MATRIX LEAST SQUARES PROBLEM $AXA^H = B$

SI-TAO LING,[1] RUI-RUI WANG,[1] and QING-BING LIU[2*]

Communicated by H. Woerdeman

Abstract. In this article, we study the perturbation analysis for the (skew) Hermitian matrix *least squares problem* (LSP). Suppose that $\mathcal{S}$ and $\widehat{\mathcal{S}}$ are two sets of solutions to the (skew) Hermitian matrix least squares problem $AXA^H = B$ and the perturbed Hermitian matrix least squares problem $\widehat{A}\widehat{X}\widehat{A}^H = \widehat{B}$, respectively. For any given $X \in \mathcal{S}$, we derive general expressions of the least squares solutions $\widehat{X} \in \widehat{\mathcal{S}}$ that are closest to X, and we present the corresponding distances between them under appropriate norms. Perturbation bounds for the nearest least squares solutions are further derived.

1. Introduction

The application of a numerical algorithm to the solution of the linear least squares problem (LSP) inevitably generates round-off errors that result from data or computer processing operations. This highlights the importance of performing perturbation analysis of LSPs (see [8]). Ding and Huang [4] presented some perturbation analysis results for least squares solutions to the operator equation $Tx = y$ in Hilbert spaces. A componentwise perturbation bound for the linear system $Ax = b$ was deduced in [14] in cases where the coefficient matrix A is rank deficient. For linear systems with multiple right-hand sides $AX = B$, the componentwise backward error and componentwise condition number under Hölder p-norms, as well as a perturbation bound for the minimum p-norm solutions, were

Received Jul. 17, 2017; Accepted Oct. 10, 2017.
First published online Apr. 4, 2018.
*Corresponding author.
2010 *Mathematics Subject Classification.* Primary 15A24; Secondary 15A09, 47A55.
Keywords. Hermitian least squares solution, perturbation bound, contraction matrix, norm-preserving dilation, nearest solution.

derived in [6], and normwise, mixed, and componentwise condition numbers were derived in [16].

In [9], the perturbation analysis for the matrix least squares problem $AXB = C$ was studied by an alternative method. According to the theory of generalized inverses and the classical norm-preserving dilation theorem (see [3]), perturbation bounds for the least squares solutions were deduced. If the coefficient matrices are required to satisfy $B = A^H$ and $C = \pm C^H$, then the perturbation analysis is converted to the (skew) Hermitian matrix LSP

$$AXA^H = B \tag{1.1}$$

with $B = \pm B^H$.

There are a number of investigations focused on analyzing and computing solutions to (1.1) with some special requirements from differing points of view. For example, symmetric, positive semidefinite, and positive definite real solutions (see [2]), antisymmetric orthosymmetric solutions (see [15]), general nonnegative definite solutions (see [7]), possible minimal rank nonnegative definite solutions (see [5]), rank-constrained Hermitian nonnegative definite solutions (see [17]), and rank-constrained least square Hermitian nonnegative definite solutions under Frobenius norm (see [13]) were considered through the generalized inverses or the *singular value decomposition* (SVD) of matrices. By applying the norm-preserving dilation theorem (see [18]) and the *Hermitian-type* (*skew Hermitian-type*) *generalized singular value decomposition* (HGSVD, SHGSVD) technique, minimum rank (skew) Hermitian solutions to the matrix approximation problem under spectral norm were investigated (see [10]). More generally, the admissible inertias and ranks of the expressions $A - BXB^H - CYC^H$ with unknowns X and Y were studied in [1].

Suppose that $A \in \mathbb{C}^{m\times n}$, $B = \pm B^H \in \mathbb{C}^{m\times m}$ are given matrices. We consider the (skew) Hermitian matrix LSP (1.1) and the perturbed (skew) Hermitian matrix least squares problem

$$\widehat{A}\widehat{X}\widehat{A}^H = \widehat{B}, \tag{1.2}$$

where $\widehat{A} = A + \Delta A, \widehat{B} = B + \Delta B$, $\Delta A, \Delta B$ are small perturbation matrices and $\widehat{B} = \pm\widehat{B}^H$. For any given least squares solution X of (1.1), we characterize general formulas of the least squares solutions $\widehat{X}$ of (1.2) that are closest to X, and present the corresponding distances between them under appropriate norms. Perturbation bounds for the nearest least squares solutions are further derived.

The rest of the article is organized as follows. In Section 2, we briefly review some results on special cases of norm-preserving dilations for further discussions. In Section 3, we study the perturbation analysis for the Hermitian matrix LSP. We first specify a general formula of the solution $\widehat{X}$ of the perturbed Hermitian matrix LSP (1.2) that is closest to any given solution X of (1.1) under appropriate norms. Applying the derived nearest solutions, we deduce the perturbation bounds for solutions of the Hermitian matrix LSP by adding some appropriate restrictions. Similar results for the skew Hermitian matrix LSP are stated in Section 4 without proofs, because the techniques are analogous to those of the Hermitian matrix LSP.

We use the following notation throughout this article:

(i) $\mathbb{C}^{m\times n}$ is the set of all $m\times n$ matrices with complex entries;
(ii) $\mathbb{C}_{\mathrm{H}}^{n\times n}$ is the set of all $n\times n$ Hermitian matrices with complex entries;
(iii) $\mathbb{C}_{\mathrm{K}}^{n\times n}$ is the set of all $n\times n$ skew-Hermitian matrices with complex entries;
(iv) $\operatorname{rank}(A)$, $\mathcal{R}(A)$, $\mathcal{N}(A)$, A^H denote rank, range, null space, and conjugate transpose of A, respectively;
(v) $A^\dagger$ is the Moore–Penrose inverse of the matrix A that uniquely exists;
(vi) $\|\cdot\|_{\mathrm{F}}$ and $\|\cdot\|_2$ represent the matrix Frobenius norm and the spectral norm, respectively; $\|\cdot\|$ stands for the matrix Frobenius norm or spectral norm.

2. Preliminaries

In this section, we list some lemmas which we will use. Let the sets of least squares solutions to the matrix LSPs (1.1) and (1.2) be $\mathcal{S}$ and $\widehat{\mathcal{S}}$, respectively. Then the least squares solutions to (1.1) are characterized by the following results.

Lemma 2.1 ([11, Lemma 1.3])**.** *Suppose that $A\in\mathbb{C}^{m\times n}$, $B\in\mathbb{C}_{\mathrm{H}}^{m\times m}$. Then the matrix equation (1.1) has a Hermitian solution if and only if $\mathcal{R}(B)\subseteq\mathcal{R}(A)$. In this case, the general Hermitian solution is*

$$X = A^\dagger B(A^\dagger)^H + (I_n - A^\dagger A)Z + Z^H(I_n - A^\dagger A)$$

for arbitrary $Z\in\mathbb{C}^{n\times n}$, among which $X_{\mathrm{LS}} = A^\dagger B(A^\dagger)^H$ is the smallest under Frobenius norm. The Hermitian solution to (1.1) is unique if and only if $A^\dagger A = I_n$, that is, $\operatorname{rank}(A) = n$.

Denote the four orthogonal projections related to a complex matrix A by

$$P_A = AA^\dagger, \qquad P_A^\perp = I_m - AA^\dagger, \qquad P_{A^H} = A^\dagger A, \qquad P_{A^H}^\perp = I_n - A^\dagger A.$$

Then the perturbation analysis on the Moore–Penrose inverse yields the following results.

Lemma 2.2 (see [12])**.** *Suppose that $A, \widehat{A}\in\mathbb{C}^{m\times n}$, $\widehat{A} = A + E$. Then*

$$\begin{aligned}\widehat{A}^\dagger - A^\dagger &= -\widehat{A}^\dagger E A^\dagger + \widehat{A}^\dagger P_A^\perp - P_{\widehat{A}^H}^\perp A^\dagger\\ &= -\widehat{A}^\dagger P_{\widehat{A}} E P_{A^H} A^\dagger + \widehat{A}^\dagger P_{\widehat{A}} P_A^\perp - P_{\widehat{A}^H}^\perp P_{A^H} A^\dagger\\ &= -\widehat{A}^\dagger P_{\widehat{A}} E P_{A^H} A^\dagger + (\widehat{A}^H\widehat{A})^\dagger E^H P_A^\perp + P_{\widehat{A}^H}^\perp E^H (AA^H)^\dagger.\end{aligned}$$

We refer to a matrix as a *contraction* if its spectral norm is less than or equal to 1. The following lemma is a special case of the Davis–Kahan–Weinberger solutions of norm-preserving dilations (see [3]).

Lemma 2.3 ([9, Lemma 2.3])**.** *For a given matrix $A\in\mathbb{C}^{m\times n}$ with $\|A\|_2 = \mu$, let*

$$Q = (\mu^2 I_n - A^H A)^{\frac{1}{2}}, \qquad Q_* = (\mu^2 I_m - AA^H)^{\frac{1}{2}}.$$

Then

(1) *there exists a matrix* $B \in \mathbb{C}^{l\times n}$ *such that*

$$\min_{B\in\mathbb{C}^{l\times n}} \left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|_2 = \mu,$$

where B *has the form* $B = KQ$ *with* $K \in \mathbb{C}^{l\times n}$ *an arbitrary contraction;*

(2) *there exists a matrix* $C \in \mathbb{C}^{m\times k}$ *such that*

$$\min_{C\in\mathbb{C}^{l\times n}} \left\| \begin{bmatrix} A, & C \end{bmatrix} \right\|_2 = \mu,$$

where C *has the form* $C = Q_*L$ *with* $L \in \mathbb{C}^{m\times k}$ *an arbitrary contraction.*

Norm-preserving dilations of Hermitian-type matrices were studied by Zheng [18], described as follows.

Lemma 2.4 ([10, Lemma 2.1])**.** *Suppose that* $A \in \mathbb{C}_{\mathrm{H}}^{m\times m}$*,* $B \in \mathbb{C}^{m\times n}$ *satisfy*

$$\left\| \begin{bmatrix} A & B \end{bmatrix} \right\|_2 = \mu.$$

Then there exists $D \in \mathbb{C}_{\mathrm{H}}^{n\times n}$ *such that*

$$\min_{D\in\mathbb{C}_{\mathrm{H}}^{n\times n}} \left\| \begin{bmatrix} A & B \\ B^H & D \end{bmatrix} \right\|_2 = \mu.$$

Moreover, a general form of D *with this property is*

$$D = -KAK^H + \mu(I_n - KK^H)^{\frac{1}{2}} Z (I_n - KK^H)^{\frac{1}{2}},$$

where

$$K^H = \left[(\mu^2 I_m - A^2)^{\frac{1}{2}}\right]^{\dagger} B,$$

and $Z \in \mathbb{C}_{\mathrm{H}}^{n\times n}$ *is an arbitrary Hermitian contraction matrix.*

As a continuation, another special case of norm-preserving dilation theorem for skew-Hermitian-type matrices is the following, which can be found in Lemma 4.1 of [10], but we modify it here slightly.

Lemma 2.5. *Suppose that* $A \in \mathbb{C}_{\mathrm{K}}^{m\times m}$*,* $B \in \mathbb{C}^{m\times n}$ *satisfy*

$$\left\| \begin{bmatrix} A & B \end{bmatrix} \right\|_2 = \mu.$$

Then there exists $D \in \mathbb{C}_{\mathrm{K}}^{n\times n}$ *such that*

$$\min_{D\in\mathbb{C}_{\mathrm{K}}^{n\times n}} \left\| \begin{bmatrix} A & B \\ -B^H & D \end{bmatrix} \right\|_2 = \mu.$$

Moreover, a general form of D *with this property is*

$$D = KAK^H + \mu(I_n - KK^H)^{\frac{1}{2}} Z (I_n - KK^H)^{\frac{1}{2}},$$

where

$$K^H = \left[(\mu^2 I_m + A^2)^{\frac{1}{2}}\right]^{\dagger} B,$$

and $Z \in \mathbb{C}_{\mathrm{K}}^{n\times n}$ *is an arbitrary skew-Hermitian contraction matrix.*

3. Perturbation analysis for the Hermitian matrix LSP

In this section, we consider the Hermitian matrix least squares problem

$$AXA^H = B, \tag{3.1}$$

where $A \in \mathbb{C}^{m\times n}$, $B = B^H \in \mathbb{C}^{m\times m}$ are given matrices. Its perturbed Hermitian matrix least squares problem is stated to be

$$\widehat{A}\widehat{X}\widehat{A}^H = \widehat{B}, \tag{3.2}$$

where $\widehat{A} = A + \Delta A, \widehat{B} = B + \Delta B$, $\Delta A, \Delta B$ are small perturbation matrices and $\widehat{B} = \widehat{B}^H$. We first specify the explicit formula of the Hermitian least squares solution of the perturbed matrix LSP (3.2) that is closest to any given Hermitian least squares solution of the matrix LSP (3.1).

Theorem 3.1. *Suppose that $A, \widehat{A} = A + \Delta A \in \mathbb{C}^{m\times n}$, $B, \widehat{B} = B + \Delta B \in \mathbb{C}_{\mathrm{H}}^{m\times m}$ are given matrices. For any given Hermitian solution X of the Hermitian matrix least squares problem (3.1), there exists a unique solution of the perturbed Hermitian matrix least squares problem (3.2) with the form*

$$\widehat{X}_m = \widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H + X - \widehat{A}^\dagger\widehat{A}X\widehat{A}^\dagger\widehat{A} \tag{3.3}$$

such that

$$\min_{\widehat{X}\in\widehat{\mathcal{S}}_{\mathrm{H}}} \|\widehat{X} - X\|_{\mathrm{F}} = \|\widehat{X}_m - X\|_{\mathrm{F}} = \left\|\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - \widehat{A}^\dagger\widehat{A}X\widehat{A}^\dagger\widehat{A}\right\|_{\mathrm{F}}, \tag{3.4}$$

where $\widehat{\mathcal{S}}_{\mathrm{H}}$ is the set of Hermitian solutions of (3.2).

Proof. For any Hermitian least squares solution $\widehat{X}$ of (3.2), from Lemma 2.1, we have the splitting

$$\begin{aligned}\widehat{X} &= \widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H + (I_n - \widehat{A}^\dagger\widehat{A})\widehat{Z} + \widehat{Z}^H(I_n - \widehat{A}^\dagger\widehat{A})\\ &= \widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H + (I_n - \widehat{A}^\dagger\widehat{A})\widehat{Z}\widehat{A}^\dagger\widehat{A} + \widehat{A}^\dagger\widehat{A}\widehat{Z}^H(I_n - \widehat{A}^\dagger\widehat{A})\\ &\quad + (I_n - \widehat{A}^\dagger\widehat{A})(\widehat{Z} + \widehat{Z}^H)(I_n - \widehat{A}^\dagger\widehat{A}),\end{aligned} \tag{3.5}$$

where $\widehat{Z} \in \mathbb{C}^{n\times n}$ is arbitrary. Therefore,

$$\begin{aligned}\widehat{X} - X &= \widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H + (I_n - \widehat{A}^\dagger\widehat{A})\widehat{Z}\widehat{A}^\dagger\widehat{A} + \widehat{A}^\dagger\widehat{A}\widehat{Z}^H(I_n - \widehat{A}^\dagger\widehat{A})\\ &\quad + (I_n - \widehat{A}^\dagger\widehat{A})(\widehat{Z} + \widehat{Z}^H)(I_n - \widehat{A}^\dagger\widehat{A}) - \big[\widehat{A}^\dagger\widehat{A}X\widehat{A}^\dagger\widehat{A} + (I_n - \widehat{A}^\dagger\widehat{A})X\widehat{A}^\dagger\widehat{A}\\ &\quad + \widehat{A}^\dagger\widehat{A}X(I_n - \widehat{A}^\dagger\widehat{A}) + (I_n - \widehat{A}^\dagger\widehat{A})X(I_n - \widehat{A}^\dagger\widehat{A})\big]\\ &= \widehat{A}^\dagger\widehat{A}\big[\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - X\big]\widehat{A}^\dagger\widehat{A} + (I_n - \widehat{A}^\dagger\widehat{A})(\widehat{Z} - X)\widehat{A}^\dagger\widehat{A}\\ &\quad + \widehat{A}^\dagger\widehat{A}(\widehat{Z}^H - X)(I_n - \widehat{A}^\dagger\widehat{A})\\ &\quad + (I_n - \widehat{A}^\dagger\widehat{A})(\widehat{Z} + \widehat{Z}^H - X)(I_n - \widehat{A}^\dagger\widehat{A}).\end{aligned} \tag{3.6}$$

Taking Frobenius norms on both sides of (3.6), we have

$$\begin{aligned}\|\widehat{X}-X\|_{\mathrm{F}}^2 &= \big\|\widehat{A}^\dagger\widehat{A}\big[\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - X\big]\widehat{A}^\dagger\widehat{A}\big\|_{\mathrm{F}}^2 + 2\big\|(I_n-\widehat{A}^\dagger\widehat{A})(\widehat{Z}-X)\widehat{A}^\dagger\widehat{A}\big\|_{\mathrm{F}}^2\\ &\quad + \big\|(I_n-\widehat{A}^\dagger\widehat{A})(\widehat{Z}+\widehat{Z}^H-X)(I_n-\widehat{A}^\dagger\widehat{A})\big\|_{\mathrm{F}}^2\\ &\geq \big\|\widehat{A}^\dagger\widehat{A}\big[\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - X\big]\widehat{A}^\dagger\widehat{A}\big\|_{\mathrm{F}}^2.\end{aligned}$$

It is clear that the last inequality becomes an equality if and only if

$$\begin{aligned}(I_n-\widehat{A}^\dagger\widehat{A})\widehat{Z}\widehat{A}^\dagger\widehat{A} &= (I_n-\widehat{A}^\dagger\widehat{A})X\widehat{A}^\dagger\widehat{A},\\ (I_n-\widehat{A}^\dagger\widehat{A})(\widehat{Z}+\widehat{Z}^H)(I_n-\widehat{A}^\dagger\widehat{A}) &= (I_n-\widehat{A}^\dagger\widehat{A})X(I_n-\widehat{A}^\dagger\widehat{A}).\end{aligned}\tag{3.7}$$

Substituting the equalities in (3.7) into (3.5), we obtain

$$\begin{aligned}\widehat{X}_m &= \widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H + (I_n-\widehat{A}^\dagger\widehat{A})X\widehat{A}^\dagger\widehat{A} + \widehat{A}^\dagger\widehat{A}X(I_n-\widehat{A}^\dagger\widehat{A})\\ &\quad + (I_n-\widehat{A}^\dagger\widehat{A})X(I_n-\widehat{A}^\dagger\widehat{A})\\ &= \widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H + X - \widehat{A}^\dagger\widehat{A}X\widehat{A}^\dagger\widehat{A},\end{aligned}$$

which is exactly the expression of $\widehat{X}_m$ in (3.3). □

From Theorem 3.1, we see that, for a given Hermitian least squares solution X of the Hermitian matrix LSP (3.1), the nearest Hermitian least squares solution $\widehat{X}_m$ is unique under the matrix Frobenius norm. However, $\widehat{X}_m$ may not be unique under spectral norm. To prove this, we first introduce the singular value decomposition (SVD) of the coefficient matrix $\widehat{A}$. Suppose that $\widehat{A}\in\mathbb{C}_r^{m\times n}$ with $\mathrm{rank}(A)=r$. Let

$$\widehat{A} = \big[\widehat{U}_1,\quad \widehat{U}_2\big]\,\mathrm{diag}\left(\widehat{\Sigma}_{\widehat{A}},\quad 0\right)\big[\widehat{V}_1,\quad \widehat{V}_2\big]^H = \widehat{U}_1\widehat{\Sigma}_{\widehat{A}}\widehat{V}_1^H \tag{3.8}$$

be the SVD of $\widehat{A}$, where $\widehat{U}_1^H\widehat{U}_1 = \widehat{V}_1^H\widehat{V}_1 = I_r$ and $\widehat{\Sigma}_{\widehat{A}} > 0$ is a diagonal matrix. Then we have the following two orthogonal projectors:

$$\widehat{A}^\dagger\widehat{A} = \widehat{V}_1\widehat{V}_1^H, \qquad I_n - \widehat{A}^\dagger\widehat{A} = \widehat{V}_2\widehat{V}_2^H.$$

Theorem 3.2. *Suppose that $A, \widehat{A} = A+\Delta A\in\mathbb{C}^{m\times n}$, $B, \widehat{B} = B+\Delta B\in\mathbb{C}_{\mathrm{H}}^{m\times m}$ are given matrices. For any given Hermitian solution X of the Hermitian matrix least squares problem (3.1), there exist Hermitian solutions $\widehat{X}_m$ of the Hermitian matrix least squares problem (3.2) such that*

$$\min_{\widehat{X}\in\widehat{\mathcal{S}}_{\mathrm{H}}}\|\widehat{X}-X\|_2 = \|\widehat{X}_m - X\|_2 = \big\|\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - \widehat{A}^\dagger\widehat{A}X\widehat{A}^\dagger\widehat{A}\big\|_2, \tag{3.9}$$

and a general form of $\widehat{X}_m$ is

$$\begin{aligned}\widehat{X}_m = \widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H &+ X - \widehat{A}^\dagger\widehat{A}X\widehat{A}^\dagger\widehat{A}\\ &- (I_n-\widehat{A}^\dagger\widehat{A})K\widehat{A}^\dagger\widehat{A}\mathcal{A}\widehat{A}^\dagger\widehat{A}K^H(I_n-\widehat{A}^\dagger\widehat{A})\\ &+ (I_n-\widehat{A}^\dagger\widehat{A})\widehat{L}^H\widehat{A}^\dagger\widehat{A}(\mu^2 I_n-\mathcal{A}^2)^{\frac{1}{2}}\widehat{A}^\dagger\widehat{A}\\ &+ \widehat{A}^\dagger\widehat{A}(\mu^2 I_n-\mathcal{A}^2)^{\frac{1}{2}}\widehat{A}^\dagger\widehat{A}\widehat{L}(I_n-\widehat{A}^\dagger\widehat{A})\end{aligned}$$

$$
\begin{aligned}
&+ \mu(I_n - \widehat{A}^{\dagger}\widehat{A})(I_n - KK^H)^{\frac{1}{2}}(I_n - \widehat{A}^{\dagger}\widehat{A}) \\
&\times \widehat{Z}(I_n - \widehat{A}^{\dagger}\widehat{A})(I_n - KK^H)^{\frac{1}{2}}(I_n - \widehat{A}^{\dagger}\widehat{A}),
\end{aligned} \tag{3.10}
$$

where $\mu = \|\mathcal{A}\|_2$ *and*

$$
\begin{aligned}
\mathcal{A} &= \widehat{A}^{\dagger}\widehat{B}(\widehat{A}^{\dagger})^H - \widehat{A}^{\dagger}\widehat{A}X\widehat{A}^{\dagger}\widehat{A}, \\
\mathcal{B} &= \widehat{A}^{\dagger}\widehat{A}(\mu^2 I_n - \mathcal{A}^2)^{\frac{1}{2}}\widehat{A}^{\dagger}\widehat{A}\widehat{L}(I_n - \widehat{A}^{\dagger}\widehat{A}), \\
K^H &= \widehat{A}^{\dagger}\widehat{A}\big[(\mu^2 I_n - \mathcal{A}^2)^{\frac{1}{2}}\big]^{\dagger}\widehat{A}^{\dagger}\widehat{A}\mathcal{B}(I_n - \widehat{A}^{\dagger}\widehat{A}),
\end{aligned}
$$

in which $\widehat{L} \in \mathbb{C}^{n\times n}$ *is an arbitrary contraction, and* $\widehat{Z} \in \mathbb{C}_{\mathrm{H}}^{n\times n}$ *is an arbitrary Hermitian contraction.*

Proof. From the expression of $\widehat{X} - X$ in (3.6) and the SVD of $\widehat{A}$ in (3.8), we have

$$
\begin{aligned}
\|\widehat{X} - X\|_2 &= \big\| \big[\widehat{V}_1, \quad \widehat{V}_2\big]^H (\widehat{X} - X) \big[\widehat{V}_1, \quad \widehat{V}_2\big] \big\|_2 \\
&= \left\| \begin{bmatrix} \mathcal{A}_1 & \mathcal{B}_1 \\ \mathcal{B}_1^H & \mathcal{D}_1 \end{bmatrix} \right\|_2,
\end{aligned}
$$

where

$$
\begin{aligned}
&\mathcal{A}_1 = \widehat{V}_1^H\big[\widehat{A}^{\dagger}\widehat{B}(\widehat{A}^{\dagger})^H - X\big]\widehat{V}_1, \qquad \mathcal{B}_1 = \widehat{V}_1^H(\widehat{Z}^H - X)\widehat{V}_2, \\
&\mathcal{D}_1 = \widehat{V}_2^H(\widehat{Z} + \widehat{Z}^H - X)\widehat{V}_2.
\end{aligned}
$$

Let

$$
\begin{aligned}
&\mathcal{A} = \widehat{V}_1\mathcal{A}_1\widehat{V}_1^H = \widehat{A}^{\dagger}\widehat{B}(\widehat{A}^{\dagger})^H - \widehat{A}^{\dagger}\widehat{A}X\widehat{A}^{\dagger}\widehat{A}, \\
&\mathcal{B} = \widehat{V}_1\mathcal{B}_1\widehat{V}_2^H, \qquad \mathcal{D} = \widehat{V}_2\mathcal{D}_1\widehat{V}_2^H,
\end{aligned} \tag{3.11}
$$

and let

$$
\hat{\mu} = \big\| \big[\mathcal{A}_1, \quad \mathcal{B}_1\big] \big\|_2.
$$

From Lemma 2.3, we observe that

$$
\begin{aligned}
\min_{\widehat{Z}\in\mathbf{C}^{n\times n}} \hat{\mu} &= \min_{\widehat{Z}\in\mathbf{C}^{n\times n}} \big\| \big[\mathcal{A}_1, \quad \mathcal{B}_1\big] \big\|_2 \\
&= \|\mathcal{A}_1\|_2 = \|\widehat{V}_1\mathcal{A}_1\widehat{V}_1^H\|_2 = \|\mathcal{A}\|_2 \\
&= \big\|\widehat{A}^{\dagger}\widehat{B}(\widehat{A}^{\dagger})^H - \widehat{A}^{\dagger}\widehat{A}X\widehat{A}^{\dagger}\widehat{A}\big\|_2 =: \mu
\end{aligned}
$$

with the choice

$$
\mathcal{B}_1 = (\mu^2 I_r - \mathcal{A}_1^2)^{\frac{1}{2}}\widetilde{L},
$$

in which $\widetilde{L} \in \mathbb{C}^{r\times(n-r)}$ is an arbitrary contraction. From (3.11) we have

$$
\begin{aligned}
\mathcal{B} = \widehat{V}_1\mathcal{B}_1\widehat{V}_2^H &= \widehat{A}^{\dagger}\widehat{A}(\widehat{Z}^H - X)(I_n - \widehat{A}^{\dagger}\widehat{A}) \\
&= \widehat{V}_1(\mu^2 I_r - \mathcal{A}_1^2)^{\frac{1}{2}}\widetilde{L}\widehat{V}_2^H \\
&= \widehat{A}^{\dagger}\widehat{A}(\mu^2 I_n - \mathcal{A}^2)^{\frac{1}{2}}\widehat{A}^{\dagger}\widehat{A}\widehat{L}(I_n - \widehat{A}^{\dagger}\widehat{A}),
\end{aligned}
$$

in which $\widehat{L} = \widehat{V}_1\widetilde{L}\widehat{V}_2^H$ is also a contraction. Thus,

$$\begin{aligned}
&\widehat{A}^\dagger\widehat{A}\widehat{Z}^H(I_n - \widehat{A}^\dagger\widehat{A})\\
&\quad= \widehat{A}^\dagger\widehat{A}X(I_n - \widehat{A}^\dagger\widehat{A}) + \widehat{A}^\dagger\widehat{A}(\mu^2 I_n - \mathcal{A}^2)^{\frac{1}{2}}\widehat{A}^\dagger\widehat{A}\widehat{L}(I_n - \widehat{A}^\dagger\widehat{A}).
\end{aligned} \tag{3.12}$$

From Lemma 2.4, we know that

$$\min_{\mathcal{D}_1}\|\widehat{X} - X\|_2 = \mu = \big\|\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - \widehat{A}^\dagger\widehat{A}X\widehat{A}^\dagger\widehat{A}\big\|_2$$

with the choice

$$\begin{aligned}
\mathcal{D}_1 &= \widehat{V}_2^H(\widehat{Z} + \widehat{Z}^H - X)\widehat{V}_2\\
&= -K_1\mathcal{A}_1K_1^H + \mu(I_{n-r} - K_1K_1^H)^{\frac{1}{2}}\widetilde{Z}(I_{n-r} - K_1K_1^H)^{\frac{1}{2}},
\end{aligned}$$

where

$$K_1^H = \big[(\mu^2 I_r - \mathcal{A}_1^2)^{\frac{1}{2}}\big]^\dagger\mathcal{B}_1 \in \mathbb{C}^{r\times(n-r)}, \tag{3.13}$$

and $\widetilde{Z} \in \mathbb{C}_{\mathrm{H}}^{(n-r)\times(n-r)}$ is an arbitrary Hermitian contraction.

It is easy to verify that

$$\begin{aligned}
&(I_n - \widehat{A}^\dagger\widehat{A})(\widehat{Z} + \widehat{Z}^H - X)(I_n - \widehat{A}^\dagger\widehat{A})\\
&\quad= \widehat{V}_2\mathcal{D}_1\widehat{V}_2^H\\
&\quad= -\widehat{V}_2K_1\mathcal{A}_1K_1^H\widehat{V}_2^H + \mu\widehat{V}_2(I_{n-r} - K_1K_1^H)^{\frac{1}{2}}\widetilde{Z}(I_{n-r} - K_1K_1^H)^{\frac{1}{2}}\widehat{V}_2^H\\
&\quad= -(I_n - \widehat{A}^\dagger\widehat{A})K\widehat{A}^\dagger\widehat{A}\mathcal{A}\widehat{A}^\dagger\widehat{A}K^H(I_n - \widehat{A}^\dagger\widehat{A})\\
&\qquad+ \mu(I_n - \widehat{A}^\dagger\widehat{A})(I_n - KK^H)^{\frac{1}{2}}(I_n - \widehat{A}^\dagger\widehat{A})\\
&\qquad\times \widehat{Z}(I_n - \widehat{A}^\dagger\widehat{A})(I_n - KK^H)^{\frac{1}{2}}(I_n - \widehat{A}^\dagger\widehat{A}),
\end{aligned}$$

where $\widehat{Z} = \widehat{V}_2\widetilde{Z}\widehat{V}_2^H \in \mathbb{C}_{\mathrm{H}}^{n\times n}$ is also an arbitrary Hermitian contraction, and

$$K = \widehat{V}_2K_1\widehat{V}_1^H \in \mathbb{C}^{n\times r}. \tag{3.14}$$

Therefore,

$$\begin{aligned}
&(I_n - \widehat{A}^\dagger\widehat{A})(\widehat{Z} + \widehat{Z}^H)(I_n - \widehat{A}^\dagger\widehat{A})\\
&\quad= (I_n - \widehat{A}^\dagger\widehat{A})X(I_n - \widehat{A}^\dagger\widehat{A}) - (I_n - \widehat{A}^\dagger\widehat{A})K\widehat{A}^\dagger\widehat{A}\mathcal{A}\widehat{A}^\dagger\widehat{A}K^H(I_n - \widehat{A}^\dagger\widehat{A})\\
&\qquad+ \mu(I_n - \widehat{A}^\dagger\widehat{A})(I_n - KK^H)^{\frac{1}{2}}(I_n - \widehat{A}^\dagger\widehat{A})\\
&\qquad\times \widehat{Z}(I_n - \widehat{A}^\dagger\widehat{A})(I_n - KK^H)^{\frac{1}{2}}(I_n - \widehat{A}^\dagger\widehat{A}).
\end{aligned} \tag{3.15}$$

Furthermore, by (3.11), (3.13), and (3.14), we have

$$\begin{aligned}
K^H &= \widehat{V}_1\big[(\mu^2 I_r - \mathcal{A}_1^2)^{\frac{1}{2}}\big]^\dagger\mathcal{B}_1\widehat{V}_2^H\\
&= \widehat{A}^\dagger\widehat{A}\big[(\mu^2 I_n - \mathcal{A}^2)^{\frac{1}{2}}\big]^\dagger\widehat{A}^\dagger\widehat{A}\mathcal{B}(I_n - \widehat{A}^\dagger\widehat{A}).
\end{aligned}$$

Combining (3.12) and (3.15) with the explicit formula of $\widehat{X}$ in (3.5), we obtain

$$\begin{aligned}\widehat{X}_m &= \widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H + (I_n - \widehat{A}^\dagger\widehat{A})X\widehat{A}^\dagger\widehat{A} + (I_n - \widehat{A}^\dagger\widehat{A})\widehat{L}^H\widehat{A}^\dagger\widehat{A}(\mu^2 I_n - \mathcal{A}^2)^{\frac{1}{2}}\widehat{A}^\dagger\widehat{A}\\ &\quad + \widehat{A}^\dagger\widehat{A}X(I_n - \widehat{A}^\dagger\widehat{A}) + \widehat{A}^\dagger\widehat{A}(\mu^2 I_n - \mathcal{A}^2)^{\frac{1}{2}}\widehat{A}^\dagger\widehat{A}\widehat{L}(I_n - \widehat{A}^\dagger\widehat{A})\\ &\quad + (I_n - \widehat{A}^\dagger\widehat{A})X(I_n - \widehat{A}^\dagger\widehat{A}) - (I_n - \widehat{A}^\dagger\widehat{A})K\widehat{A}^\dagger\widehat{A}\mathcal{A}\widehat{A}^\dagger\widehat{A}K^H(I_n - \widehat{A}^\dagger\widehat{A})\\ &\quad + \mu(I_n - \widehat{A}^\dagger\widehat{A})(I_n - KK^H)^{\frac{1}{2}}(I_n - \widehat{A}^\dagger\widehat{A})\\ &\quad \times \widehat{Z}(I_n - \widehat{A}^\dagger\widehat{A})(I_n - KK^H)^{\frac{1}{2}}(I_n - \widehat{A}^\dagger\widehat{A})\\ &= \widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H + X - \widehat{A}^\dagger\widehat{A}X\widehat{A}^\dagger\widehat{A} - (I_n - \widehat{A}^\dagger\widehat{A})K\widehat{A}^\dagger\widehat{A}\mathcal{A}\widehat{A}^\dagger\widehat{A}K^H(I_n - \widehat{A}^\dagger\widehat{A})\\ &\quad + (I_n - \widehat{A}^\dagger\widehat{A})\widehat{L}^H\widehat{A}^\dagger\widehat{A}(\mu^2 I_n - \mathcal{A}^2)^{\frac{1}{2}}\widehat{A}^\dagger\widehat{A}\\ &\quad + \widehat{A}^\dagger\widehat{A}(\mu^2 I_n - \mathcal{A}^2)^{\frac{1}{2}}\widehat{A}^\dagger\widehat{A}\widehat{L}(I_n - \widehat{A}^\dagger\widehat{A})\\ &\quad + \mu(I_n - \widehat{A}^\dagger\widehat{A})(I_n - KK^H)^{\frac{1}{2}}(I_n - \widehat{A}^\dagger\widehat{A})\\ &\quad \times \widehat{Z}(I_n - \widehat{A}^\dagger\widehat{A})(I_n - KK^H)^{\frac{1}{2}}(I_n - \widehat{A}^\dagger\widehat{A}).\end{aligned}$$

This completes the proof. □

Note that the Hermitian least squares solution $\widehat{X}_m$ of the form in (3.3) is a special case of (3.10) with $\widehat{L} = 0$ and $\widehat{Z} = 0$. To derive perturbation bounds for the Hermitian least squares solutions of the Hermitian matrix LSP (3.1), we assume that perturbations for the coefficient matrix A keep rank invariant. Using the results obtained in Theorems 3.1–3.2, we now present the following main results of this section.

Theorem 3.3. *Suppose that $A, \widehat{A} = A + \Delta A \in \mathbb{C}^{m\times n}$, $B, \widehat{B} = B + \Delta B \in \mathbb{C}_{\mathrm{H}}^{m\times m}$ are given matrices with* $\mathrm{rank}(A) = \mathrm{rank}(\widehat{A})$ *and* $\|A^\dagger\|_2\|\Delta A\|_2 < 1$. *Then*

(1) *for the Hermitian least squares solution $X_{\mathrm{LS}} = A^\dagger B(A^\dagger)^H$ with the minimum Frobenius norm, we have the estimate*

$$\begin{aligned}&\|\widehat{X}_m - X_{\mathrm{LS}}\|\\ &\quad = \big\|\widehat{A}^\dagger\widehat{A}\big[\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - A^\dagger B(A^\dagger)^H\big]\widehat{A}^\dagger\widehat{A}\big\|\\ &\quad \le 2\|A^\dagger\|_2\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta A P_{A^H}\|_2 + 2\|A^\dagger\|_2^2\|A^\dagger B\|\|P_A^\perp \Delta A P_{A^H}\|_2\\ &\qquad + \|A^\dagger\|_2\|A^\dagger\|\|P_A\Delta B P_A\|_2 + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big),\end{aligned} \tag{3.16}$$

where $\|\cdot\|$ is either the matrix Frobenius norm or spectral norm;

(2) *for any given Hermitian least squares solution $X \in \mathcal{S}_{\mathrm{H}}$ of the form*

$$X = A^\dagger B(A^\dagger)^H + (I_n - A^\dagger A)Z + Z^H(I_n - A^\dagger A) \tag{3.17}$$

with $Z \in \mathbb{C}^{n\times n}$, let $\widehat{X}_m$ be the same as in Theorem 3.1 or Theorem 3.2. Then we have the estimate

$$\begin{aligned}&\|\widehat{X}_m - X\|\\ &\quad \le 2\|A^\dagger\|_2\Big(\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta A P_{A^H}\|_2 + \|A^\dagger\|_2\|A^\dagger B\|\|P_A^\perp \Delta A P_{A^H}\|_2\end{aligned}$$

$$+ \frac{1}{2}\|A^\dagger\|\|P_A\Delta B P_A\|_2 + \|P_{A^H}^\perp Z P_{A^H}\|\|P_A \Delta A P_{A^H}^\perp\|_2\Big) + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big), \tag{3.18}$$

where $\|\cdot\|$ is either the matrix Frobenius norm or spectral norm.

Proof. (1) For $\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - A^\dagger B(A^\dagger)^H$, we have the following decomposition:

$$\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - A^\dagger B(A^\dagger)^H = (\widehat{A}^\dagger - A^\dagger)\widehat{B}(\widehat{A}^\dagger)^H + A^\dagger(\widehat{B}-B)(\widehat{A}^\dagger)^H + A^\dagger B(\widehat{A}^\dagger - A^\dagger)^H.$$

Furthermore, from Lemma 2.2, we have

$$\begin{aligned}
&\|\widehat{X}_m - X_{\mathrm{LS}}\| \\
&= \big\|\widehat{A}^\dagger\widehat{A}\big[\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - A^\dagger B(A^\dagger)^H\big]\widehat{A}^\dagger\widehat{A}\big\| \\
&= \big\|-\widehat{A}^\dagger\Delta A A^\dagger\widehat{B}(\widehat{A}^\dagger)^H + \widehat{A}^\dagger(\widehat{A}^\dagger)^H(\Delta A)^H(I_m - AA^\dagger)\widehat{B}(\widehat{A}^\dagger)^H \\
&\quad + \widehat{A}^\dagger\widehat{A}A^\dagger\Delta B(\widehat{A}^\dagger)^H \\
&\quad - \widehat{A}^\dagger\widehat{A}A^\dagger B(A^\dagger)^H(\Delta A)^H(\widehat{A}^\dagger)^H + \widehat{A}^\dagger\widehat{A}A^\dagger B(I_m - AA^\dagger)\Delta A\widehat{A}^\dagger(\widehat{A}^\dagger)^H\big\| \\
&\le \|\widehat{A}^\dagger\|_2\big\|A^\dagger\widehat{B}(\widehat{A}^\dagger)^H\big\|\|P_{\widehat{A}}\Delta A P_{A^H}\|_2 + \|\widehat{A}^\dagger\|_2^2\big\|\widehat{B}(\widehat{A}^\dagger)^H\big\|\|P_A^\perp\Delta A P_{\widehat{A}^H}\|_2 \\
&\quad + \|A^\dagger\|_2\|\widehat{A}^\dagger\|\|P_A\Delta B P_{\widehat{A}}\|_2 + \|\widehat{A}^\dagger\|_2\big\|A^\dagger B(A^\dagger)^H\big\|\|P_{\widehat{A}}\Delta A P_{A^H}\|_2 \\
&\quad + \|\widehat{A}^\dagger\|_2^2\|A^\dagger B\|\|P_A^\perp\Delta A P_{\widehat{A}^H}\|_2 \\
&= 2\|A^\dagger\|_2\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta A P_{A^H}\|_2 + 2\|A^\dagger\|_2^2\|A^\dagger B\|\|P_A^\perp\Delta A P_{A^H}\|_2 \\
&\quad + \|A^\dagger\|_2\|A^\dagger\|\|P_A\Delta B P_A\|_2 + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big),
\end{aligned}$$

where $\|\cdot\|$ is either the matrix Frobenius norm or spectral norm. In the last equality, we have used the fact that

$$\frac{\|A^\dagger\|_2}{1+\|A^\dagger\|_2\|\Delta A\|_2} \le \|\widehat{A}^\dagger\|_2 \le \frac{\|A^\dagger\|_2}{1-\|A^\dagger\|_2\|\Delta A\|_2}, \tag{3.19}$$

which holds by applying the conditions of the theorem. Then we yield the estimate in (3.16).

(2) By applying the assertion (3.16), we have from (3.17), (3.4), and (3.9) that

$$\begin{aligned}
&\|\widehat{X}_m - X\| \\
&= \big\|\widehat{A}^\dagger\widehat{A}(\widehat{A}^\dagger)^H - \widehat{A}^\dagger\widehat{A}X\widehat{A}(\widehat{A}^\dagger)^H\big\| \\
&= \big\|\widehat{A}^\dagger\widehat{A}\big[\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - A^\dagger B(A^\dagger)^H - (I_n - A^\dagger A)Z - Z^H(I_n - A^\dagger A)\big]\widehat{A}^\dagger\widehat{A}\big\| \\
&= \big\|\widehat{A}^\dagger\widehat{A}\big[\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - A^\dagger B(A^\dagger)^H\big]\widehat{A}^\dagger\widehat{A} \\
&\quad - \widehat{A}^\dagger\Delta A(I_n - A^\dagger A)Z\widehat{A}^\dagger\widehat{A} - \widehat{A}^\dagger\widehat{A}Z^H(I_n - A^\dagger A)(\Delta A)^H(\widehat{A}^\dagger)^H\big\| \\
&\le 2\|A^\dagger\|_2\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta A P_{A^H}\|_2 + 2\|A^\dagger\|_2^2\|A^\dagger B\|\|P_A^\perp\Delta A P_{A^H}\|_2 \\
&\quad + \|A^\dagger\|_2\|A^\dagger\|\|P_A\Delta B P_A\|_2 + 2\|\widehat{A}^\dagger\|_2\big\|(I_n - A^\dagger A)Z\widehat{A}^\dagger\widehat{A}\big\|\|P_{\widehat{A}}\Delta A P_{A^H}^\perp\|_2 \\
&\quad + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big)
\end{aligned}$$

$$
\begin{aligned}
&= 2\|A^\dagger\|_2\Big(\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta AP_{A^H}\|_2 + \|A^\dagger\|_2\|A^\dagger B\|\|P_A^\perp \Delta AP_{A^H}\|_2 \\
&\quad + \frac{1}{2}\|A^\dagger\|\|P_A\Delta BP_A\|_2 + \|P_{A^H}^\perp ZP_{A^H}\|\|P_A\Delta AP_{A^H}^\perp\|_2\Big) \\
&\quad + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big).
\end{aligned}
$$

Note that in the last equality, we used the inequalities in (3.19) once again. Hence the estimate in (3.18) follows. □

If the Hermitian linear system (3.1) is consistent, we have the following conclusion.

Theorem 3.4. *Under the assumptions of Theorem 3.3, furthermore suppose that the Hermitian linear system (3.1) is consistent. Then*

(1) *for the minimum Frobenius norm Hermitian solution $X_{\rm LS} = \widehat{A}^\dagger \widehat{B}(\widehat{A}^\dagger)^H$, we have the estimate*

$$
\begin{aligned}
&\|\widehat{X}_m - X_{\rm LS}\| \\
&\quad = \big\|\widehat{A}^\dagger \widehat{A}\big[\widehat{A}^\dagger \widehat{B}(\widehat{A}^\dagger)^H - A^\dagger B(A^\dagger)^H\big]\widehat{A}^\dagger \widehat{A}\big\| \\
&\quad \le 2\|A^\dagger\|_2\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta AP_{A^H}\|_2 + \|A^\dagger\|_2\|A^\dagger\|\|P_A\Delta BP_A\|_2 \\
&\qquad + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big),
\end{aligned} \tag{3.20}
$$

where $\|\cdot\|$ is either the matrix Frobenius norm or spectral norm;

(2) *for any given Hermitian solution $X \in \mathcal{S}_{\rm H}$ of the form (3.17), let $\widehat{X}_m$ be as in Theorem 3.1 or Theorem 3.2. Then we have the estimate*

$$
\begin{aligned}
&\|\widehat{X}_m - X\| \\
&\quad \le 2\|A^\dagger\|_2\big(\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta AP_{A^H}\|_2 + \|P_{A^H}^\perp ZP_{A^H}\|\|P_A\Delta AP_{A^H}^\perp\|_2\big) \\
&\qquad + \|A^\dagger\|_2\|A^\dagger\|\|P_A\Delta BP_A\|_2 + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big),
\end{aligned} \tag{3.21}
$$

where $\|\cdot\|$ is either the matrix Frobenius norm or spectral norm.

Proof. (1) The consistency of the linear system (3.1) implies that $B = AA^\dagger BAA^\dagger$. Therefore,

$$
\begin{aligned}
&\|\widehat{X}_m - X_{\rm LS}\| \\
&\quad = \big\|\widehat{A}^\dagger \widehat{A}\big[\widehat{A}^\dagger \widehat{B}(\widehat{A}^\dagger)^H - A^\dagger B(A^\dagger)^H\big]\widehat{A}^\dagger \widehat{A}\big\| \\
&\quad = \big\|-\widehat{A}^\dagger \Delta AA^\dagger \widehat{B}(\widehat{A}^\dagger)^H + \widehat{A}^\dagger(\widehat{A}^\dagger)^H(\Delta A)^H(I_m - AA^\dagger)\widehat{B}(\widehat{A}^\dagger)^H \\
&\qquad + \widehat{A}^\dagger \widehat{A}A^\dagger \Delta B(\widehat{A}^\dagger)^H \\
&\qquad - \widehat{A}^\dagger \widehat{A}A^\dagger B(A^\dagger)^H(\Delta A)^H(\widehat{A}^\dagger)^H + \widehat{A}^\dagger \widehat{A}A^\dagger B(I_m - AA^\dagger)\Delta A\widehat{A}^\dagger(\widehat{A}^\dagger)^H\big\| \\
&\quad = \big\|-\widehat{A}^\dagger \Delta AA^\dagger \widehat{B}(\widehat{A}^\dagger)^H + \widehat{A}^\dagger(\widehat{A}^\dagger)^H(\Delta A)^H(I_m - AA^\dagger)\Delta B(\widehat{A}^\dagger)^H \\
&\qquad + \widehat{A}^\dagger \widehat{A}A^\dagger \Delta B(\widehat{A}^\dagger)^H - \widehat{A}^\dagger \widehat{A}A^\dagger B(A^\dagger)^H(\Delta A)^H(\widehat{A}^\dagger)^H\big\| \\
&\quad \le 2\|A^\dagger\|_2\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta AP_{A^H}\|_2 + \|A^\dagger\|_2\|A^\dagger\|\|P_A\Delta BP_A\|_2 \\
&\qquad + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big).
\end{aligned}
$$

Note that in the last inequality, we used the inequalities in (3.19) to get the estimate in (3.20).

(2) By applying the assertion in (3.20), we have from (3.4) and (3.9) that

$$
\begin{aligned}
&\|\widehat{X}_m - X\| \\
&\quad = \big\|\widehat{A}^\dagger\widehat{A}(\widehat{A}^\dagger)^H - \widehat{A}^\dagger\widehat{A}X\widehat{A}(\widehat{A}^\dagger)^H\big\| \\
&\quad = \big\|\widehat{A}^\dagger\widehat{A}\big[\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - A^\dagger B(A^\dagger)^H - (I_n - A^\dagger A)Z - Z^H(I_n - A^\dagger A)\big]\widehat{A}^\dagger\widehat{A}\big\| \\
&\quad = \big\|\widehat{A}^\dagger\widehat{A}\big[\widehat{A}^\dagger\widehat{B}(\widehat{A}^\dagger)^H - A^\dagger B(A^\dagger)^H\big]\widehat{A}^\dagger\widehat{A} \\
&\qquad - \widehat{A}^\dagger\Delta A(I_n - A^\dagger A)Z\widehat{A}^\dagger\widehat{A} - \widehat{A}^\dagger\widehat{A}Z^H(I_n - A^\dagger A)(\Delta A)^H(\widehat{A}^\dagger)^H\big\| \\
&\quad \le 2\|A^\dagger\|_2\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta AP_{A^H}\|_2 + \|A^\dagger\|_2\|A^\dagger\|\|P_A\Delta BP_A\|_2 \\
&\qquad + 2\|\widehat{A}^\dagger\|_2\big\|(I_n - A^\dagger A)Z\widehat{A}^\dagger\widehat{A}\big\|\|P_{\widehat{A}}\Delta AP_{A^H}^\perp\|_2 + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big) \\
&\quad = 2\|A^\dagger\|_2\big(\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta AP_{A^H}\|_2 + \|P_{A^H}^\perp ZP_{A^H}\|\|P_A\Delta AP_{A^H}^\perp\|_2\big) \\
&\qquad + \|A^\dagger\|_2\|A^\dagger\|\|P_A\Delta BP_A\|_2 + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big).
\end{aligned}
$$

Note that in the last equality, we used the inequalities in (3.19) once again. Hence the estimate in (3.21) follows. □

4. Perturbation analysis for the skew-Hermitian matrix LSP

Since the perturbation analysis for the skew-Hermitian matrix least squares problem can be treated using the same technique as in the preceding section, we only state the main results without proofs here.

For the skew-Hermitian matrix least squares problem

$$AXA^H = B, \tag{4.1}$$

where $A \in \mathbb{C}^{m\times n}$, $B = -B^H \in \mathbb{C}^{m\times m}$ are given matrices, its perturbed Hermitian matrix least squares problem is described to be

$$\widehat{A}\widehat{X}\widehat{A}^H = \widehat{B}, \tag{4.2}$$

where $\widehat{A} = A + \Delta A$, $\widehat{B} = B + \Delta B$, $\Delta A, \Delta B$ are small perturbation matrices and $\widehat{B} = -\widehat{B}^H$. The general solutions of (4.1) are given by the following lemma.

Lemma 4.1 ([11, Lemma 1.5]). *Suppose that $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}_{\mathrm{K}}^{m\times m}$. Then the matrix equation (4.1) has a skew-Hermitian solution if and only if $\mathcal{R}(B) \subseteq \mathcal{R}(A)$. In this case, the general skew-Hermitian solution is*

$$X = A^\dagger B(A^\dagger)^H + (I_n - A^\dagger A)Z - Z^H(I_n - A^\dagger A)$$

for arbitrary $Z \in \mathbb{C}^{n\times n}$. The skew-Hermitian solution to (4.1) is unique if and only if $A^\dagger A = I_n$; that is, $\operatorname{rank}(A) = n$.

According to Lemma 4.1 and applying the same proof as Theorem 3.1, we have the following result.

Theorem 4.2. *Suppose that* $A, \widehat{A} = A + \Delta A \in \mathbb{C}^{m\times n}$, $B, \widehat{B} = B + \Delta B \in \mathbb{C}_{\mathrm{K}}^{m\times m}$ *are given matrices. For any given skew-Hermitian solution* X *of the skew-Hermitian matrix LSP (4.1) there exists a unique solution of the perturbed skew-Hermitian matrix LSP (4.2) with the form*

$$\widehat{X}_m = \widehat{A}^\dagger \widehat{B}(\widehat{A}^\dagger)^H + X - \widehat{A}^\dagger \widehat{A} X \widehat{A}^\dagger \widehat{A} \tag{4.3}$$

such that

$$\min_{\widehat{X}\in\widehat{\mathcal{S}}_{\mathrm{K}}} \|\widehat{X} - X\|_{\mathrm{F}} = \|\widehat{X}_m - X\|_{\mathrm{F}} = \left\|\widehat{A}^\dagger \widehat{B}(\widehat{A}^\dagger)^H - \widehat{A}^\dagger \widehat{A} X \widehat{A}^\dagger \widehat{A}\right\|_{\mathrm{F}},$$

where $\widehat{\mathcal{S}}_{\mathrm{K}}$ *is the set of skew-Hermitian solutions of (4.2).*

If the distances are measured by the spectral norm, according to Lemmas 2.5 and 4.1, we have the following result.

Theorem 4.3. *Suppose that* $A, \widehat{A} = A + \Delta A \in \mathbb{C}^{m\times n}$, $B, \widehat{B} = B + \Delta B \in \mathbb{C}_{\mathrm{K}}^{m\times m}$ *are given matrices. For any given skew-Hermitian solution* X *of the skew-Hermitian matrix LSP (4.1) there exist skew-Hermitian solutions* $\widehat{X}_m$ *of the skew-Hermitian matrix LSP (4.2) such that*

$$\min_{\widehat{X}\in\widehat{\mathcal{S}}_{\mathrm{K}}} \|\widehat{X} - X\|_2 = \|\widehat{X}_m - X\|_2 = \left\|\widehat{A}^\dagger \widehat{B}(\widehat{A}^\dagger)^H - \widehat{A}^\dagger \widehat{A} X \widehat{A}^\dagger \widehat{A}\right\|_2,$$

and a general form of $\widehat{X}_m$ *is*

$$\begin{aligned}\widehat{X}_m = {}& \widehat{A}^\dagger \widehat{B}(\widehat{A}^\dagger)^H + X - \widehat{A}^\dagger \widehat{A} X \widehat{A}^\dagger \widehat{A} + (I_n - \widehat{A}^\dagger \widehat{A}) K \widehat{A}^\dagger \widehat{A} \mathcal{A} \widehat{A}^\dagger \widehat{A} K^H (I_n - \widehat{A}^\dagger \widehat{A}) \\ & - (I_n - \widehat{A}^\dagger \widehat{A}) \widehat{L}^H \widehat{A}^\dagger \widehat{A} (\mu^2 I_n + \mathcal{A}^2)^{\frac{1}{2}} \widehat{A}^\dagger \widehat{A} \\ & + \widehat{A}^\dagger \widehat{A} (\mu^2 I_n + \mathcal{A}^2)^{\frac{1}{2}} \widehat{A}^\dagger \widehat{A} \widehat{L} (I_n - \widehat{A}^\dagger \widehat{A}) \\ & + \mu (I_n - \widehat{A}^\dagger \widehat{A})(I_n - KK^H)^{\frac{1}{2}} (I_n - \widehat{A}^\dagger \widehat{A}) \\ & \times \widehat{Z} (I_n - \widehat{A}^\dagger \widehat{A})(I_n - KK^H)^{\frac{1}{2}} (I_n - \widehat{A}^\dagger \widehat{A}),\end{aligned} \tag{4.4}$$

where $\mu = \|\mathcal{A}\|_2$ *and*

$$\begin{aligned}\mathcal{A} &= \widehat{A}^\dagger \widehat{B}(\widehat{A}^\dagger)^H - \widehat{A}^\dagger \widehat{A} X \widehat{A}^\dagger \widehat{A}, \\ \mathcal{B} &= \widehat{A}^\dagger \widehat{A} (\mu^2 I_n + \mathcal{A}^2)^{\frac{1}{2}} \widehat{A}^\dagger \widehat{A} \widehat{L} (I_n - \widehat{A}^\dagger \widehat{A}), \\ K^H &= \widehat{A}^\dagger \widehat{A} \big[(\mu^2 I_n + \mathcal{A}^2)^{\frac{1}{2}}\big]^\dagger \widehat{A}^\dagger \widehat{A} \mathcal{B} (I_n - \widehat{A}^\dagger \widehat{A}),\end{aligned}$$

in which $\widehat{L} \in \mathbb{C}^{n\times n}$ *is an arbitrary contraction and* $\widehat{Z} \in \mathbb{C}_{\mathrm{K}}^{n\times n}$ *is an arbitrary skew-Hermitian contraction.*

Analogous to the case of the Hermitian LSP, from Theorem 4.2 we see that, for a given skew-Hermitian least squares solution X of the skew-Hermitian matrix LSP (4.1), the nearest skew-Hermitian least squares solution $\widehat{X}_m$ is unique under the matrix Frobenius norm. However, $\widehat{X}_m$ may not be unique under spectral norm. Moreover, the skew-Hermitian least squares solution $\widehat{X}_m$ of the form in (4.3) is a special case of (4.4) with $\widehat{L} = 0$ and $\widehat{Z} = 0$.

Perturbation bounds for the skew-Hermitian least square solutions of (4.1) are presented according to the results derived in Theorems 4.2–4.3, by adding a restriction that the perturbations of the coefficient matrix A keep rank-invariant.

Theorem 4.4. *Suppose that* $A, \widehat{A} = A + \Delta A \in \mathbb{C}^{m\times n}$, $B, \widehat{B} = B + \Delta B \in \mathbb{C}_{\mathrm{K}}^{m\times m}$ *are given matrices with* $\operatorname{rank}(A) = \operatorname{rank}(\widehat{A})$ *and* $\|A^\dagger\|_2\|\Delta A\|_2 < 1$. *Then*

(1) *for the skew-Hermitian least squares solution* $X_{\mathrm{LS}} = A^\dagger B(A^\dagger)^H$ *with the minimum Frobenius norm, we have the estimate*

$$\begin{aligned}
&\|\widehat{X}_m - X_{\mathrm{LS}}\| \\
&\quad = \big\|\widehat{A}^\dagger \widehat{A}\big[\widehat{A}^\dagger \widehat{B}(\widehat{A}^\dagger)^H - A^\dagger B(A^\dagger)^H\big]\widehat{A}^\dagger \widehat{A}\big\| \\
&\quad \le 2\|A^\dagger\|_2\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta A P_{A^H}\|_2 + 2\|A^\dagger\|_2^2\|A^\dagger B\|\|P_A^\perp \Delta A P_{A^H}\|_2 \\
&\qquad + \|A^\dagger\|_2\|A^\dagger\|\|P_A\Delta B P_A\|_2 + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big),
\end{aligned}$$

where $\|\cdot\|$ *is either the matrix Frobenius norm or spectral norm;*

(2) *for any given skew-Hermitian least squares solution* $X \in \mathcal{S}_{\mathrm{K}}$ *of the form*

$$X = A^\dagger B(A^\dagger)^H + (I_n - A^\dagger A)Z - Z^H(I_n - A^\dagger A) \tag{4.5}$$

with $Z \in \mathbb{C}^{n\times n}$, *let* $\widehat{X}_m$ *be the same as in Theorem 4.2 or Theorem 4.3. Then we have the estimate*

$$\begin{aligned}
\|\widehat{X}_m - X\| \le{} & 2\|A^\dagger\|_2\Big(\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta A P_{A^H}\|_2 + \|A^\dagger\|_2\|A^\dagger B\|\|P_A^\perp \Delta A P_{A^H}\|_2 \\
& + \frac{1}{2}\|A^\dagger\|\|P_A\Delta B P_A\|_2 + \|P_{A^H}^\perp Z P_{A^H}\|\|P_A \Delta A P_{A^H}^\perp\|_2\Big) \\
& + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big),
\end{aligned}$$

where $\|\cdot\|$ *is either the matrix Frobenius norm or spectral norm.*

If the skew-Hermitian linear system (4.1) is consistent, we have the following conclusion.

Theorem 4.5. *Under the assumptions of Theorem 4.4, furthermore suppose that the skew-Hermitian linear system (4.1) is consistent. Then*

(1) *for the minimum Frobenius norm skew-Hermitian solution* $X_{\mathrm{LS}} = \widehat{A}^\dagger \widehat{B}(\widehat{A}^\dagger)^H$, *we have the estimate*

$$\begin{aligned}
\|\widehat{X}_m - X_{\mathrm{LS}}\| &= \big\|\widehat{A}^\dagger \widehat{A}\big[\widehat{A}^\dagger \widehat{B}(\widehat{A}^\dagger)^H - A^\dagger B(A^\dagger)^H\big]\widehat{A}^\dagger \widehat{A}\big\| \\
&\le 2\|A^\dagger\|_2\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta A P_{A^H}\|_2 + \|A^\dagger\|_2\|A^\dagger\|\|P_A\Delta B P_A\|_2 \\
&\quad + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big),
\end{aligned}$$

where $\|\cdot\|$ *is either the matrix Frobenius norm or spectral norm;*

(2) *for any given skew-Hermitian solution* $X \in \mathcal{S}_{\mathrm{K}}$ *of the form (4.5), let* $\widehat{X}_m$ *be as in Theorem 4.2 or Theorem 4.3. Then we have the estimate*

$$\begin{aligned}
\|\widehat{X}_m - X\| \le{} & 2\|A^\dagger\|_2\big(\big\|A^\dagger B(A^\dagger)^H\big\|\|P_A\Delta A P_{A^H}\|_2 + \|P_{A^H}^\perp Z P_{A^H}\|\|P_A\Delta A P_{A^H}^\perp\|_2\big) \\
& + \|A^\dagger\|_2\|A^\dagger\|\|P_A\Delta B P_A\|_2 + \mathcal{O}\big(\|\Delta A\|_2^2 + \|\Delta B\|_2^2\big),
\end{aligned}$$

where $\|\cdot\|$ *is either the matrix Frobenius norm or spectral norm.*

Acknowledgments. Ling's work was partially supported by Fundamental Research Funds for the Central Universities grant 2015QNA47.

References

1. D. Chu, Y. S. Hung, and H. J. Woerdeman, *Inertia and rank characterizations of some matrix expressions*, SIAM J. Matrix Anal. Appl. **31** (2009), no. 3, 1187–1226. Zbl 1198.15010. MR2558819. DOI 10.1137/080712945. 436
2. H. Dai and P. Lancaster, *Linear matrix equations from an inverse problem of vibration theory*, Linear Algebra Appl. **246** (1996), 31–47. Zbl 0861.15014. MR1407657. DOI 10.1016/0024-3795(94)00311-4. 436
3. C. Davis, W. M. Kahan, and H. F. Weinberger, *Norm-preserving dilations and their applications to optimal error bounds*, SIAM J. Numer. Anal. **19** (1982), no. 3, 445–469. Zbl 0491.47003. MR0656462. DOI 10.1137/0719029. 436, 437
4. J. Ding and L. J. Huang, *On the perturbation of the least squares solutions in Hilbert spaces*, Linear Algebra Appl. **212/213** (1994), 487–500. Zbl 0817.47017. MR1306994. DOI 10.1016/0024-3795(94)90418-9. 435
5. J. Gross, *Nonnegative-definite and positive-definite solutions to the matrix equation $AXA^* = B$ revisited*, Linear Algebra Appl. **321** (2000), no. 1—3, 123–129. Zbl 0984.15011. MR1799988. DOI 10.1016/S0024-3795(00)00033-1. 436
6. D. J. Higham and N. J. Higham, *Componentwise perturbation theory for linear systems with multiple right-hand sides*, Linear Algebra Appl. **174** (1992), 111–129. Zbl 0801.65043. MR1176455. DOI 10.1016/0024-3795(92)90046-D. 436
7. C. G. Khatri and S. K. Mitra, *Hermitian and nonnegative definite solutions of linear matrix equations*, SIAM J. Appl. Math. **31** (1976), no. 4, 579–585. Zbl 0359.65033. MR0417212. DOI 10.1137/0131050. 436
8. M. Konstantinov, D. Gu, V. Mehrmann, and P. Petkov, *Perturbation Theory for Matrix Equations*, Stud. Comput. Math. **9**, North-Holland, Amsterdam, 2003. Zbl 1025.15017. MR1991778. 435
9. S. Ling, M. Wei, and Z. Jia, *Perturbation analysis for the matrix least squares problem $AXB = C$*, J. Comput. Appl. Math. **273** (2015), 150–159. Zbl 1295.65044. MR3239243. DOI 10.1016/j.cam.2014.06.007. 436, 437
10. D. Shen, M. Wei, and Y. Liu, *Minimum rank (skew) Hermitian solutions to the matrix approximation problem in the spectral norm*, J. Comput. Appl. Math. **288** (2015), 351–365. Zbl 1320.65065. MR3349628. DOI 10.1016/j.cam.2015.04.033. 436, 438
11. Y. Tian and Y. Liu, *Extremal ranks of some symmetric matrix expressions with applications*, SIAM J. Matrix Anal. Appl. **28** (2006), no. 3, 890–905. Zbl 1123.15001. MR2262987. DOI 10.1137/S0895479802415545. 437, 446
12. P.-Å. Wedin, *Perturbation theory for pseudo-inverses*, BIT **13** (1973), 217–232. Zbl 0263.65047. MR0336982. 437
13. M. Wei and Q. Wang, *On rank-constrained Hermitian nonnegative-definite least squares solutions to the matrix equation $AXA^H = B$*, Inter. J. Comput. Math. **84** (2007), no. 6, 945–952. Zbl 1129.15012. MR2335377. DOI 10.1080/00207160701458344. 436
14. Y. Wei, Y. Cao, and H. Xiang, *A note on the componentwise perturbation bounds of matrix inverse and linear systems*, Appl. Math. Comput. **169** (2005), no. 2, 1221–1236. Zbl 1083.65048. MR2174716. DOI 10.1016/j.amc.2004.10.065. 435
15. Q. Xiao, X. Hu, and L. Zhang, *The anti-symmetric ortho-symmetric solutions of the matrix equation $A^TXA = D$*, Electron. J. Linear Algebra **18** (2009), 21–29. Zbl 1175.65047. MR2471193. DOI 10.13001/1081-3810.1291. 436
16. H. Zhang, H. Xiang, and Y. Wei, *Condition numbers for linear systems and Kronecker product linear systems with multiple right-hand sides*, Int. J. Comput. Math. **84** (2007), no. 12, 1805–1817. Zbl 1134.15004. MR2377161. DOI 10.1080/00207160701332390. 436

17. X. Zhang and M. Cheng, *The rank-constrained Hermitian nonnegative-definite and positive-definite solutions to the matrix equation $AXA^* = B$*, Linear Algebra Appl. **370** (2003), 163–174. Zbl 1026.15011. MR1994325. DOI 10.1016/S0024-3795(03)00385-9. 436
18. D. Zheng, *Further study and generalization of Kahan's matrix extension theorem*, SIAM J. Matrix Anal. Appl. **17** (1996), no. 3, 621–631. Zbl 0855.47004. MR1397249. DOI 10.1137/0617037. 436, 438

[1]School of Mathematics, China University of Mining and Technology, Xuzhou Jiangsu 221116, People's Republic of China.

E-mail address: lingsitao2004@163.com; doublerui612@gmail.com

[2]Business School, Zhejiang Wanli University, Ningbo Zhejiang 315100, People's Republic of China.

E-mail address: toby1997@163.com