Multidimensional Total Least Squares Problem with Linear Equality Constraints *

Qiaohua Liu†, Zhigang Jia‡, and Yimin Wei§

Abstract. Many recent data analysis models are mathematically characterized by a multidimensional total least squares problem with linear equality constraint (TLSE). In this paper, an explicit solution is firstly derived for the multidimensional TLSE problem, as well as the solvability conditions. With applying the perturbation theory of invariant subspace, the multidimensional TLSE problem is proved equivalent to a multidimensional unconstrained weighed total least squares problem in the limit sense. The Kronecker-product-based formulae are also given for the normwise, mixed and componentwise condition numbers of the multidimensional TLSE solution of minimum Frobenius norm, and their computable upper bounds are also provided to reduce the storage and computational cost. All these results are appropriate for the single right-hand-side case and the multidimensional total least squares problem, which are two especial cases of the multidimensional TLSE problem. In numerical experiments, the multidimensional TLSE model is successfully applied to the color image deblurring and denoising for the first time, and the numerical results also indicate the effectiveness of the condition numbers.

Key words. multidimensional TLSE problem; weighted TLS problem; multidimensional TLS problem; condition number.

AMS subject classifications. 65F35, 65F20

1. Introduction.

The multidimensional total least square (TLS) model, which arises in many data fitting and estimation problems, finds a "best" fit to the overdetermined system $Ax \approx B$, where $A \in \mathbb{R}^{q\times n}$ $(q > n)$ and $B \in \mathbb{R}^{q\times d}$ are contaminated by some noise. It determines perturbations E to the coefficient matrix A and F to the matrix B measured by the Frobenius norm such that

$$\min_{E,F} \|[E \quad F]\|_F, \quad \text{subject to} \quad (A+E)X = B+F. \tag{1.1}$$

After the minimizer $[\hat{E} \quad \hat{F}]$ is found, a solution X to the consistent corrected system $(A+\hat{E})X = B+\hat{F}$ is called the TLS solution. The TLS model, was originally proposed in 1901 for data fitting problem [33], but has not caught much attention for a long time. In 1980, Golub and Van Loan [14] introduced this model into the numerical linear algebra area. Since then, it has been attracting more and more attention and now the TLS model is applied in a broad class of scientific disciplines such as system identification [21], image processing [31, 32], speech and audio processing [17, 22], etc. An overview of applications, theory, and computational methods of the TLS problem, we refer to [14, 18, 29, 41, 48, 51].

*This paper is supported in part by the National Natural Science Foundation of China under grants 12090011, 11771188, 11771099 and 12171210; the Priority Academic Program Development Project (PAPD); the Top-notch Academic Programs Project (No. PPZY2015A013) of Jiangsu Higher Education Institutions; Innovation Program of Shanghai Municipal Education Commission.

†Department of Mathematics, Shanghai University, Shanghai 200444, P.R. China. Email: qhliu@shu.edu.cn

‡1. School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou 221116, P.R. China; 2. The Research Institute of Mathematical Science, Jiangsu Normal University, Xuzhou 221116, China. Email: zhgjia@jsnu.edu.cn

§Corresponding author. School of Mathematical Sciences and Key Laboratory of Mathematics for Nonlinear Sciences, Fudan University, Shanghai, 200433, P.R. China. Email: ymwei@fudan.edu.cn

An extension of TLS model is the following multidimensional TLS problem with equality constraints (TLSE):

$$\min_{E,F} \|[E \quad F]\|_F, \quad \text{subject to} \quad (A+E)X = B+F, \quad CX = D, \tag{1.2}$$

where the matrix $C \in \mathbb{R}^{p\times n}$ is of full row rank and the right-hand-side matrix D has d columns. Within our knowledge, this problem has never been studied in the literature. We will present the solvability conditions and an explicit solution of (1.2), and define several new condition numbers to make sensitivity analysis.

The TLSE with a single right-hand side was first presented by Dowling, Degroat and Linebarger [10] in 1992. A well-known application of the single right-hand-side TLSE is the linear prediction [35] method for solving the frequency estimation of the signal model

$$y_k = \sum_{i=1}^{M} a_i \mathrm{e}^{\mathrm{j}2\pi f_i k} + \omega_k, \quad k = 0, 1, 2, \cdots, N-1, \ \mathrm{j}^2 = -1,$$

where $\{y_k\}$ and $\{\omega_k\}$ are the measured samples and additive zero mean white Gaussian noise samples, respectively, N is the data length, f_i and a_i are the frequency and amplitude of the ith sinusoid. A linear prediction equation $Ax \approx b$ as

$$\begin{bmatrix} y_0 & y_1 & \cdots & y_{L-1} \\ y_1 & y_2 & \cdots & y_L \\ \vdots & \vdots & \ddots & \vdots \\ y_{N-L-1} & y_{N-L} & \cdots & y_{N-2} \end{bmatrix} \begin{bmatrix} c_L \\ c_{L-1} \\ \vdots \\ c_1 \end{bmatrix} \approx \begin{bmatrix} y_L \\ y_{L+1} \\ \vdots \\ y_{N-1} \end{bmatrix},$$

is solved, where L is the prediction filter order satisfying $M \le L \le N-2M$, and after the linear prediction vector c is evaluated, the quantity $\mathrm{e}^{\mathrm{j}2\pi f_i}$ can be determined from the zeros of the characteristic polynomial $z^L - c_1 z^{L-1} - \ldots - c_{L-1}z - c_L = 0$. If some frequencies, say $f_1, \ldots, f_p$ are already known, then it gives the constraint equation $[C \quad d]\begin{bmatrix} x \\ -1 \end{bmatrix} = 0$ with $[1 \ \mathrm{e}^{\mathrm{j}2\pi f_i} \ \mathrm{e}^{\mathrm{j}4\pi f_i} \ \ldots \ \mathrm{e}^{\mathrm{j}2\pi f_i L}]$ as the i-th ($1 \le i \le p$) row of $[C \quad d]$.

Conventional algorithm for single right-hand-side TLSE is on the basis of QR and singular value decomposition (SVD) matrix factorizations [10]. Further investigations on the TLSE problem were performed in [26, 37]. In [26], the uniqueness condition of the TLSE solution was analyzed and proved to be an approximation of an unconstrained weighted TLS problem in the limit sense. This observation stimulated a QR-based inverse iteration method, which is more efficient than the iterative algorithm in [37].

As a vital definition in numerical analysis, the condition number measures the worst-case sensitivity of the solution of a problem to small perturbations in the input data. When C, D are zero matrices and $d = 1$, the multidimensional TLSE problem becomes the standard TLS problem with a single right-hand side, whose first order perturbation analysis and condition numbers have been widely studied in [1, 6, 8, 9, 13, 19, 23, 28, 47, 49]. By making use of the perturbation results in [1, 19, 23] and the close relation of the TLSE to an unconstrained weighted TLS problem, Liu and Jia [25] derived closed formulae for condition numbers of the TLSE problem. Further investigations on the perturbation results of TLSE problem with different-magnitude input data were given in [24], where the results unify the ones for the least squares problem with equality constraint (LSE) studied by [4, 7, 45]. On the other hand, when C, D are zero matrices and $d > 1$, the TLSE problem becomes the multidimensional TLS problem. In 1992, Wei [43, 44] provided perturbation analysis for TLS with more than one solution. Recently, Zheng, Meng and Wei [30, 50] made further investigations on the condition numbers for the TLS problem with unique and multiple solutions, respectively.

In this paper, we aim to derive an explicit solution for the multidimensional TLSE problem and present the general formulae of its condition numbers and their computable upper bounds, which have not been addressed in the literature. With the invariant subspace perturbation theorem, we prove that the multidimensional TLSE problem is equivalent to a multidimensional weighted TLS problem, with a large weight assigned on the constraint; and thereby any direct and iterative algorithms for standard TLS are feasible for TLSE based on the weighting method. The multidimensional TLSE model is successfully applied to the color image deblurring and denoising problem for the first time, and the numerical results indicate the effectiveness of our method.

Throughout this paper, $\|\cdot\|_2$ denotes the Euclidean vector or matrix norm, I_n, 0_n, $0_{m\times n}$ denote the $n\times n$ identity matrix, $n\times n$ zero matrix, and $m\times n$ zero matrix, respectively. If subscripts are ignored, the sizes of identity and zero matrices are clear from the context. For a matrix $M\in\mathbb{R}^{m\times n}$, M^T, $M^\dagger$, $\mathcal{R}(M)$, $\sigma_i(M)(\sigma_{\min}(M))$ and $\|M\|_{\max}$ denote the transpose, the Moore-Penrose inverse, the column range space, the i-th largest (the smallest) singular value and the maximal absolute value of elements of M, respectively. $\mathrm{vec}(M)$ is an operator, which stacks the columns of M one underneath the other. The Kronecker product of $A\in\mathbb{R}^{m\times n}$ and $B\in\mathbb{R}^{s\times t}$ is defined by $A\otimes B = [a_{ij}B]$ and has the property [16, 20] below:

$$
\begin{aligned}
&\mathrm{vec}(AXB) = (B^T\otimes A)\mathrm{vec}(X), \quad (A\otimes B)(C\otimes D) = (AC)\otimes(BD),\\
&(A\otimes B)^T = A^T\otimes B^T, \quad (A\otimes B)^\dagger = A^\dagger\otimes B^\dagger, \quad \|A\otimes B\|_2 = \|A\|_2\|B\|_2,\\
&\mathrm{vec}(A^T) = \Pi_{(m,n)}\mathrm{vec}(A), \quad \Pi_{(s,m)}(A\otimes B) = (B\otimes A)\Pi_{(t,n)},
\end{aligned}
$$

where $X\in\mathbb{R}^{n\times s}$, $C\in\mathbb{R}^{n\times k}$, $D\in\mathbb{R}^{t\times r}$ and $\Pi_{(m,n)}$ is an $mn\times mn$ vec-permutation matrix.

2. Preliminaries.

In this section we first recall some well known results of the multidimensional TLS problem and then derive the solvability conditions and the explicit form for the multidimensional TLSE solution. We start with a key lemma which will be used in the proofs of our main results.

Lemma 2.1 ([50]). *Let* $Q = \begin{bmatrix} Q_{11} & Q_{12}\\ Q_{21} & Q_{22}\end{bmatrix}$ *be an n-by-n orthogonal matrix with a 2-by-2 partitioning, then*

(a) Q_{11} *has full column (row) rank if and only if* Q_{22} *has full row (column) rank;*

(b) $\|Q_{11}^\dagger\|_2 = \|Q_{22}^\dagger\|_2, \quad Q_{11}^{\dagger T} = Q_{11} - Q_{12}Q_{22}^\dagger Q_{21}, \quad Q_{11}^{\dagger T}Q_{21}^T = -Q_{12}Q_{22}^\dagger.$

2.1. The multidimensional TLS problems.

Let $L\in\mathbb{R}^{m\times n}$, $H\in\mathbb{R}^{m\times d}$ $(m\geq n+d)$, the multidimensional TLS problem is defined by

$$
\min_{E,F}\|[E \quad F]\|_F, \qquad \text{subject to} \qquad (L+E)X = H+F. \tag{2.1}
$$

Following [14], the multidimensional TLS problem (2.1) may have no solutions. In order to broad its scope of applications, the generic and nongeneric conditions for TLS solutions were further studied by Van Huffel and Vandewalle [40, 42], and then refined and generalized by Wei [44] to make the multidimensional TLS problem (2.1) meaningful in any situation.

SVD is a useful tool to characterize the TLS solution. If the skinny SVD [15, Chapter 2.4] of $[L \quad H]$ is given by

$$[L \quad H] = U\Sigma V^T, \quad \Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_{n+d}) \in \mathbb{R}^{(n+d)\times(n+d)}, \tag{2.2}$$

where $\sigma_i = \sigma_i([L \quad H])$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{n+d} \geq 0$, $U \in \mathbb{R}^{m\times(n+d)}$ and $V \in \mathbb{R}^{(n+d)\times(n+d)}$ have orthonormal columns. For an integer $t \in [0, n]$, partition

$$V = \begin{array}{c} n \\ d \\ \\ \end{array} \begin{array}{c} \begin{bmatrix} V_{11}(t) & V_{12}(t) \\ V_{21}(t) & V_{22}(t) \end{bmatrix} \\ \begin{array}{cc} t \qquad & \qquad n+d-t \end{array} \end{array}.$$

For simplicity, we denote $V_{ij} = V_{ij}(t)$ for $i, j = 1, 2$. If $\sigma_t > \sigma_{t+1}$ and $\mathrm{rank}(V_{22}) = d$ hold simultaneously [44], then a solution to the consistent linear system $\widehat{L}X = \widehat{H}$ is defined as a TLS solution to the linear approximation equation $LX \approx H$, where $\widehat{L} = U_1\Sigma_1 V_{11}^T$ and $\widehat{H} = U_1\Sigma_1 V_{21}^T$ with U_1, V_1 being, respectively, the first t columns of $U = [U_1 \quad U_2]$ and $V = [V_1 \quad V_2]$, the diagonal matrices $\Sigma_1 = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_t)$ and $\Sigma_2 = \mathrm{diag}(\sigma_{t+1}, \sigma_{t+2}, \cdots, \sigma_{n+d})$. Among all TLS solutions, the solution of minimum Frobenius norm to the compatible system is given by $X_t = -V_{12}V_{22}^{\dagger}$. By Lemma 2.1(b), it can also be expressed as

$$X_t = -V_{12}V_{22}^{\dagger} = V_{11}^{\dagger T} V_{21}^T. \tag{2.3}$$

Specially, when $\Sigma_2 = \sigma_{t+1} I_{n+d-t}$, X_t satisfies $[L \ H]^T [L \ H] \left[\begin{smallmatrix} X_t \\ -I_d \end{smallmatrix}\right] = \sigma_{t+1}^2 \left[\begin{smallmatrix} X_t \\ -I_d \end{smallmatrix}\right]$, and hence X_t minimizes the problem $\min\limits_X \frac{\|H - LX\|_F^2}{d + \|X\|_F^2}$.

Zheng, Meng and Wei [50] defined the mapping ϕ: $\mathbb{R}^{m(n+d)} \to \mathbb{R}^{nd}$ by $\phi(c) = \mathrm{vec}(X_t)$ for $c = \mathrm{vec}([L \quad H])$ and provided the first order perturbation analysis of $\phi(c)$ as

$$\begin{aligned} \mathrm{vec}(\Delta X_t) &= \phi'(c)\mathrm{vec}([\Delta L \quad \Delta H]) + \mathcal{O}(\|\Delta L\|_F^2 + \|\Delta H\|_F^2) \\ &= (H_1 + H_2) D Z \mathrm{vec}([\Delta L \quad \Delta H]) + \mathcal{O}(\|\Delta L\|_F^2 + \|\Delta H\|_F^2), \end{aligned} \tag{2.4}$$

where

$$\begin{aligned} &H_1 = \Big((V_{22}V_{22}^T)^{-1} V_{21} \otimes (V_{12} F_{V_{22}}) \Big), \qquad H_2 = \Big(V_{22}^{\dagger T} \otimes V_{11}^{\dagger T} \Big) \Pi_{(n+d-t,t)}, \\ &D = (\Sigma_1^2 \otimes I_{n+d-t} - I_t \otimes (\Sigma_2^T \Sigma_2))^{-1} \big[I_t \otimes \Sigma_2^T \qquad \Sigma_1 \otimes I_{n+d-t} \big], \quad Z = \begin{bmatrix} V_1^T \otimes U_2^T \\ \Pi_{(t,n+d-t)}(V_2^T \otimes U_1^T) \end{bmatrix}, \end{aligned}$$

$\Pi_{(n+d-t,t)}$ is a vec-permutation matrix, and $F_{V_{22}} = I - V_{22}^{\dagger} V_{22}$. From this result, the absolute normwise condition number was derived for the TLS problem in [50].

2.2. Solvability conditions and explicit solution of multi-dimensional TLSE problem. For the multidimensional TLSE problem (1.2), denote $\widetilde{A} = [A \;\; B], \widetilde{C} = [C \;\; D]$, and assume that the QR factorization of $\widetilde{C}^T$ takes the form:

$$\widetilde{C}^T = \widetilde{Q} \begin{bmatrix} \widetilde{R}_1 \\ 0 \end{bmatrix}, \qquad \widetilde{Q} = [\widetilde{Q}_1 \quad \widetilde{Q}_2], \tag{2.5}$$

in which $\widetilde{Q}_1 \in \mathbb{R}^{(n+d)\times p}$, $\widetilde{Q}_2 \in \mathbb{R}^{(n+d)\times(n+d-p)}$. Let the skinny SVD of $\widetilde{A}\widetilde{Q}_2$ be

$$\widetilde{A}\widetilde{Q}_2 = \widetilde{U}\widetilde{\Sigma}\widetilde{V}^T = [\widetilde{U}_1 \quad \widetilde{U}_2] \begin{bmatrix} \widetilde{\Sigma}_1 & 0 \\ 0 & \widetilde{\Sigma}_2 \end{bmatrix} [\widetilde{V}_1 \quad \widetilde{V}_2]^T, \tag{2.6}$$

where $\widetilde{U} \in \mathbb{R}^{q\times(n+d-p)}, \widetilde{V} \in \mathbb{R}^{(n+d-p)\times(n+d-p)}$, and the matrices $\widetilde{U}_1$, $\widetilde{V}_1$ are the submatrices of $\widetilde{U}$ and $\widetilde{V}$ by taking their first k columns, respectively. Here $0 \le k \le n-p$ is an integer such that the singular values of $\widetilde{A}\widetilde{Q}_2$ satisfy

$$\mathcal{C}(k): \qquad \widetilde{\sigma}_1 \ge \widetilde{\sigma}_2 \ge \ldots \ge \widetilde{\sigma}_k > \widetilde{\sigma}_{k+1} \ge \ldots \ge \widetilde{\sigma}_{n+d-p}. \tag{2.7}$$

Denote the diagonal matrices $\widetilde{\Sigma}_1 = \mathrm{diag}(\widetilde{\sigma}_1, \widetilde{\sigma}_2, \ldots, \widetilde{\sigma}_k)$, $\widetilde{\Sigma}_2 = \mathrm{diag}(\widetilde{\sigma}_{k+1}, \widetilde{\sigma}_{k+2}, \ldots, \widetilde{\sigma}_{n+d-p})$. In the following theorem, we give the solvability conditions and explicit form of the solution to the multidimensional TLSE problem.

Theorem 2.2. *With the notations in (2.5)–(2.7), let $t = p+k$ and $\overline{V} = \widetilde{Q}_2\widetilde{V}$ have the partition*

$$\overline{V} = \widetilde{Q}_2\widetilde{V} = \underset{\hphantom{n+d-t}k\qquad n+d-t}{[\overline{V}_1 \quad \overline{V}_2]} = \begin{array}{c} n \\ d \\ \; \end{array} \underset{\hphantom{..}k\qquad\quad n+d-t}{\begin{bmatrix} \overline{V}_{11} & \overline{V}_{12} \\ \overline{V}_{21} & \overline{V}_{22} \end{bmatrix}}. \tag{2.8}$$

If for $k = n-p$, the condition $\mathcal{C}(k)$ holds and $\overline{V}_{22}$ is nonsingular, then the unique solution of the multidimensional TLSE problem (1.2) is determined by $X_n = -\overline{V}_{12}\overline{V}_{22}^{-1}$, which is also the solution to the consistent linear system

$$\widehat{A}X = \widehat{B}, \quad \textit{subject to} \quad CX = D, \tag{2.9}$$

where

$$\widehat{A} = \widetilde{U}_1\widetilde{\Sigma}_1\overline{V}_{11}^T, \qquad \widehat{B} = \widetilde{U}_1\widetilde{\Sigma}_1\overline{V}_{21}^T. \tag{2.10}$$

Proof. Let $\widetilde{X} = \begin{bmatrix} X^T & -I_d \end{bmatrix}^T$. Notice that the constraint $CX = D$ requires $\widetilde{C}\widetilde{X} = 0$, therefore the column range $\mathcal{R}(\widetilde{X})$ of $\widetilde{X}$ lies in the null space of $\widetilde{C}$ spanned by $\widetilde{Q}_2$. Denoting $\widetilde{X} = \widetilde{Q}_2 Z$ and writing $\widetilde{A} = \widetilde{A}\widetilde{Q}_1\widetilde{Q}_1^T + \widetilde{A}\widetilde{Q}_2\widetilde{Q}_2^T$, $\widetilde{E} = [E \quad F]$, (1.2) becomes

$$\min \|[\widetilde{E}\widetilde{Q}_1 \quad \widetilde{E}\widetilde{Q}_2]\|_F, \qquad \text{s.t.} \qquad (\widetilde{A}\widetilde{Q}_2 + \widetilde{E}\widetilde{Q}_2)Z = 0, \tag{2.11}$$

where the restriction only imposed on $\widetilde{E}\widetilde{Q}_2$ means that we can choose optimal $\widetilde{E}_*$ such that $\widetilde{E}_*\widetilde{Q}_1 = 0$ and $\widetilde{A}\widetilde{Q}_2 + \widetilde{E}_*\widetilde{Q}_2$ has a null space with dimension no less than d.

Note that the condition $\widetilde{E}_*\widetilde{Q}_1 = 0$ means there exists a matrix Y such that $\widetilde{E}_* = Y\widetilde{Q}_2^T$, and (2.11) becomes

$$\min_{\operatorname{rank}(\widetilde{A}\widetilde{Q}_2+Y)\le n-p} \|Y\|_F, \quad \text{s.t.} \quad (\widetilde{A}\widetilde{Q}_2 + Y)Z = 0.$$

According to (2.6) and the well-known Eckart-Young theorem [15, Theorem 2.4.8] for the best rank-$(n-p)$ matrix approximation, the optimal Y_* satisfies $Y_* = -\widetilde{U}_2\widetilde{\Sigma}_2\widetilde{V}_2^T$, and for the optimal error matrix $\widetilde{E}_* = Y_*\widetilde{Q}_2$, the corrected system becomes

$$(\widetilde{A}\widetilde{Q}_2 - \widetilde{U}_2\widetilde{\Sigma}_2\widetilde{V}_2^T)Z = 0, \quad \textit{or} \quad \widetilde{U}_1\widetilde{\Sigma}_1(\widetilde{Q}_2\widetilde{V}_1)^T\widetilde{X} = 0.$$

Recalling that $\mathcal{R}(\widetilde{X}) \subseteq \mathcal{R}(\widetilde{Q}_2)$, $\mathcal{R}(\widetilde{X})$ lies in the range of $\overline{V}_2 = \widetilde{Q}_2\widetilde{V}_2$, i.e., there exists a $d \times d$ matrix G such that

$$\begin{bmatrix} X \\ -I_d \end{bmatrix} = \begin{bmatrix} \overline{V}_{12} \\ \overline{V}_{22} \end{bmatrix} G, \tag{2.12}$$

from which we obtain $G = -\overline{V}_{22}^{-1}$ and the unique solution is given by $X_n = -\overline{V}_{12}\overline{V}_{22}^{-1}$. ■

Remark 2.3. If for $k = n-p$, $\overline{V}_{22}$ is singular, then X_n is not well defined. In this circumstance, we can always seek an integer $0 \le k \le n-p$ such that

$$\mathcal{C}(k) \text{ holds and } \overline{V}_{22} \text{ has full row-rank,} \tag{2.13}$$

say for $k = 0$, notice that the matrix

$$\begin{bmatrix} \widetilde{Q}_1 & \widetilde{Q}_2\widetilde{V} \end{bmatrix} = \begin{array}{c} \left[\begin{array}{c|c} \widetilde{Q}_{11} & \overline{V}_{1:} \\ \hline \widetilde{Q}_{21} & \overline{V}_{2:} \end{array}\right] \end{array} \begin{array}{c} n \\ d \end{array}$$

is an orthogonal matrix, where $\widetilde{Q}_{11}$ has full column-rank, since by (2.5), $C^T = \widetilde{Q}_{11}\widetilde{R}_1$ is a full column-rank matrix. In view of Lemma 2.1(a), $\overline{V}_{2:}$ has full row-rank and a TLSE solution for the linear system (2.9)-(2.10) can be defined to make the TLSE problem meaningful in any situation. Likewise, (2.12) still holds where $G \in \mathbb{R}^{(n+d-t)\times d}$ satisfies $\overline{V}_{22}G = -I_d$. It follows that $G = -\overline{V}_{22}^{\dagger} + PK$ for an arbitrary $(n+d-t)\times d$ matrix K and $P = I_{n+d-t} - \overline{V}_{22}^{\dagger}\overline{V}_{22}$. Therefore any TLSE solution X has the form

$$X = -\overline{V}_{12}\overline{V}_{22}^{\dagger} + \overline{V}_{12}PK,$$

in which

$$\left(\overline{V}_{12}\overline{V}_{22}^{\dagger}\right)^T\overline{V}_{12}P = \overline{V}_{22}^{\dagger T}\overline{V}_{12}^T\overline{V}_{12}P = \overline{V}_{22}^{\dagger T}(I - \overline{V}_{22}^T\overline{V}_{22})P = 0,$$

and $X_t = -\overline{V}_{12}\overline{V}_{22}^{\dagger}$ is the minimum Frobenius norm solution among all TLSE solutions.

The explicit form of the solution X_t includes the one for the special case $d = 1$, where the TLSE problem is described by $Ax \approx \mathsf{b}$ subject to $Cx = \mathsf{d}$. By setting $x_C = C^{\dagger}\mathsf{d}$ and

$r_C = Ax_C - \mathsf{b}$, in [26] Liu et al. proved that if the orthonormal basis of null space of $\widetilde{C}$ is chosen as

$$\widetilde{Q}_2 = \begin{bmatrix} Q_2 & \beta^{-1}x_C \\ 0 & -\beta^{-1} \end{bmatrix}, \quad \beta = \left(1 + \|x_C\|_2^2\right)^{1/2}, \tag{2.14}$$

in which Q_2 is the orthonormal basis of the null space of C, then under the condition

$$\sigma_{n-p}(AQ_2) > \sigma_{n-p+1}([AQ_2 \quad \beta^{-1}r_C]) = \sigma_{n-p+1}(\widetilde{A}\widetilde{Q}_2) = \widetilde{\sigma}_{n-p+1}, \tag{2.15}$$

the TLSE solution x_n is unique and takes another closed-form

$$x_n = x_C - \mathcal{K}A^T r_C, \quad for \quad \mathcal{K} = Q_2(Q_2^T A^T A Q_2 - \widetilde{\sigma}_{n-p+1}^2 I_{n-p})^{-1} Q_2^T. \tag{2.16}$$

3. Close relation of TLSE to an unconstrained weighted TLS problem. In this section, we aim to interpret the solution of the multidimensional TLSE problem as an approximation of the solution to a multidimensional weighted TLS (WTLS) problem, by assigning a large weight on the constraint.

Firstly, we need to generalize Stewart's result [38] about the asymptotic behavior for the scaled SVD of $X_\epsilon = [X_1 \quad \epsilon X_2]$, based on the following perturbation theorem for invariant subspaces.

Lemma 3.1 (Chapter V, Theorem 2.7 in [39]). *Let $[Z_1 \quad Y_2] \in \mathbb{R}^{n\times n}$ be an orthogonal matrix and $\mathcal{R}(Z_1)$ is a k-dimensional simple invariant subspace of $n \times n$ matrix C such that*

$$[Z_1 \quad Y_2]^T C [Z_1 \quad Y_2] = \begin{bmatrix} L_1 & H \\ 0 & L_2 \end{bmatrix},$$

where L_1 and L_2 have no common eigenvalues, and $Y_2^T C Z_1 = 0$. (Here $\mathcal{R}(Z_1), \mathcal{R}(Y_2)$ are called the simple right and left invariant subspace of C, respectively). Given a perturbation E, let

$$[Z_1 \quad Y_2]^T E [Z_1 \quad Y_2] = \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix}.$$

Then for sufficiently small perturbation $\|E\|_2$, there exists a unique matrix P such that the columns of

$$\tilde{Z}_1 = (Z_1 + Y_2 P)(I + P^T P)^{-\frac{1}{2}}, \qquad \tilde{Y}_2 = (Y_2 - Z_1 P^T)(I + PP^T)^{-\frac{1}{2}} \tag{3.1}$$

form orthonormal bases for simple right and left invariant subspaces of $\tilde{C} = C + E$. The representations of $\tilde{C}$ with respect to $\tilde{Z}_1$, $\tilde{Y}_2$ are given by $\tilde{C}\tilde{Z}_1 = \tilde{Z}_1\tilde{L}_1$, $\tilde{C}\tilde{Y}_2 = \tilde{Y}_2\tilde{L}_2$ for

$$\begin{aligned} \tilde{L}_1 &= (I + P^T P)^{\frac{1}{2}}[L_1 + E_{11} + (H + E_{12})P](I + P^T P)^{-\frac{1}{2}}, \\ \tilde{L}_2 &= (I + PP^T)^{-\frac{1}{2}}[L_2 + E_{22} - P(H + E_{12})](I + PP^T)^{\frac{1}{2}}. \end{aligned} \tag{3.2}$$

With Lemma 3.1, the asymptotic results for the scaled SVD of X_ϵ are given as follows.

Lemma 3.2. *Let $\epsilon > 0$ be a small parameter, $X = [X_1 \quad X_2] \in \mathbb{R}^{m\times n}$ with $X_1 \in \mathbb{R}^{m\times k}$ being of full column-rank. Denote $X_\epsilon = [X_1 \quad \epsilon X_2]$, $\overline{X}_2 = X_2 - X_1 B$ with $B = X_1^\dagger X_2$. Let the skinny SVD of X_1 be $X_1 = U_1 S_1 V_1^T$, and $\overline{X}_2 = \overline{U}_2\overline{S}_2\overline{V}_2^T$, $X_\epsilon = U_\epsilon S_\epsilon V_\epsilon^T$ are SVDs of $\overline{X}_2$ and X_ϵ, respectively, then*

$$
\begin{aligned}
&S_\epsilon = \mathrm{diag}(S_1 + \mathcal{O}(\epsilon^2), \epsilon\overline{S}_2 + \mathcal{O}(\epsilon^3)),\\
&U_\epsilon = \begin{bmatrix} U_1 + \mathcal{O}(\epsilon^2) & \overline{U}_2 + \mathcal{O}(\epsilon^2)\end{bmatrix}, \quad V_\epsilon = \begin{bmatrix} V_1 + \mathcal{O}(\epsilon^2) & -\epsilon B\overline{V}_2 + \mathcal{O}(\epsilon^3) \\ \epsilon B^T V_1 + \mathcal{O}(\epsilon^3) & \overline{V}_2 + \mathcal{O}(\epsilon^2)\end{bmatrix}.
\end{aligned}
$$

Proof. Let $G = [X_1 \quad 0_{m\times(n-k)}]^T[X_1 \quad 0_{m\times(n-k)}]$ and let

$$
G_\epsilon = X_\epsilon^T X_\epsilon = G + \begin{bmatrix} 0 & \epsilon X_1^T X_2 \\ \epsilon X_2^T X_1 & \epsilon^2 X_2^T X_2 \end{bmatrix} =: G + E
$$

be the perturbed version of G. Notice that

$$
[Z_1 \quad Y_2] \quad = \quad \begin{array}{c c} & \begin{array}{cc} \end{array} \\ \begin{array}{c} k \\ n-k \end{array} & \begin{bmatrix} V_1 & 0 \\ 0 & \overline{V}_2 \end{bmatrix} \\ & \begin{array}{cc} k & n-k \end{array} \end{array}
$$

has orthonormal columns forming orthonormal basis of simple invariant subspace of G such that the representations of G with respect to Z_1, Y_2 are

$$
GZ_1 = Z_1 L_1, \qquad GY_2 = Y_2 L_2, \qquad \text{for} \quad L_1 = S_1^T S_1, \quad L_2 = 0_{n-k}. \tag{3.3}
$$

By Lemma 3.1, there exists an $(n-k)\times k$ matrix P such that $\tilde{Z}_1, \tilde{Y}_2$ with structure (3.1) form the orthonormal bases of right and left invariant subspaces of G_ϵ, respectively. Substituting (3.1) into the relation $\tilde{Y}_2^T G_\epsilon \tilde{Z}_1 = 0$, one can derive that $(Y_2 - Z_1 P^T)^T (G+E)(Z_1 + Y_2 P) = 0$. Using (3.3) and the expressions for Z_1, Y_2, we obtain

$$
P(S_1^T S_1) = Y_2^T E Z_1 - P Z_1^T E Z_1 + Y_2^T E Y_2 P - P Z_1^T E Y_2 P,
$$

from which

$$
P = \epsilon \overline{V}_2^T (X_1^\dagger X_2)^T V_1 + \mathcal{O}(\epsilon^3).
$$

From (3.1)-(3.2), $[\tilde{Z}_1 \quad \tilde{Y}_2]$ has the following form

$$
[\tilde{Z}_1 \quad \tilde{Y}_2] = \begin{bmatrix} V_1 + \mathcal{O}(\epsilon^2) & -\epsilon B\overline{V}_2 + \mathcal{O}(\epsilon^3) \\ \epsilon B^T V_1 + \mathcal{O}(\epsilon^3) & \overline{V}_2 + \mathcal{O}(\epsilon^2)\end{bmatrix},
$$

and the representations of G_ϵ with respect to $\tilde{Z}_1, \tilde{Y}_2$ are given by $\tilde{L}_1 = S_1^T S_1 + \mathcal{O}(\epsilon^2)$ and

$$
\begin{aligned}
\tilde{L}_2 &= (Y_2^T E Y_2 - P Z_1^T E Y_2)(1 + \mathcal{O}(\epsilon^2)) \\
&= \epsilon^2 \overline{V}_2^T [X_2^T X_2 - (X_1^\dagger X_2)^T (X_1^T X_2)]\overline{V}_2 + \mathcal{O}(\epsilon^4) \\
&= \epsilon^2 \overline{V}_2^T \overline{X}_2^T \overline{X}_2 \overline{V}_2 + \mathcal{O}(\epsilon^4) = \epsilon^2 \overline{S}_2^T \overline{S}_2 + \mathcal{O}(\epsilon^4).
\end{aligned}
$$

Note that $G_\epsilon = X_\epsilon^T X_\epsilon$ is symmetric and G_ϵ has $\tilde{Z}_1$ and $\tilde{Y}_2$ as the bases of its simple right and left invariant subspaces such that $\tilde{Y}_2^T G_\epsilon \tilde{Z}_1 = 0$. Therefore $\tilde{H} := [\tilde{Z}_1 \quad \tilde{Y}_2]$ satisfies

$$\tilde{H}^T X_\epsilon^T X_\epsilon \tilde{H} = \begin{bmatrix} S_1^T S_1 + \mathcal{O}(\epsilon^2) & 0 \\ 0 & \epsilon^2 \overline{S}_2^T \overline{S}_2 + \mathcal{O}(\epsilon^4) \end{bmatrix},$$

where the orthonormal columns of $\tilde{H}$ span the right singular subspace of X_ϵ, with diagonal entries of $S_1 + \mathcal{O}(\epsilon^2), \epsilon\overline{S}_2 + \mathcal{O}(\epsilon^3)$ as its singular values. By taking $V_\epsilon = \tilde{H}$ and using the relation

$$X_\epsilon V_\epsilon = [X_1 V_1 + \mathcal{O}(\epsilon^2) \quad \epsilon\overline{X}_2 \overline{V}_2 + \mathcal{O}(\epsilon^3)] = [U_1 S_1 + \mathcal{O}(\epsilon^2) \quad \epsilon\overline{U}_2 \overline{S}_2 + \mathcal{O}(\epsilon^3)],$$

we conclude that the left singular matrix U_ϵ of X_ϵ satisfies

$$U_\epsilon = \left[U_1 + \mathcal{O}(\epsilon^2) \quad \overline{U}_2 + \mathcal{O}(\epsilon^2)\right].$$

The proof is then completed. ∎

Based on Lemma 3.2, the close relation of the multidimensional TLSE problem to an unconstrained WTLS problem is illustrated below.

Theorem 3.3. *For the multidimensional TLSE problem (1.2) with the notations in (2.5)–(2.8), assume that $\overline{V}_{22}$ has full row rank, and the minimum Frobenius norm solution $X_t = -\overline{V}_{12}\overline{V}_{22}^\dagger$. Denote*

$$L_\epsilon = W_\epsilon^{-1} L = \begin{bmatrix} \epsilon^{-1} C \\ A \end{bmatrix}, \qquad H_\epsilon = W_\epsilon^{-1} H = \begin{bmatrix} \epsilon^{-1} D \\ B \end{bmatrix}, \tag{3.4}$$

where $W_\epsilon = \mathrm{diag}(\epsilon I_p, I_q)$ with ϵ being a small positive parameter. Let the multidimensional weighted TLS problem be given by

$$\min_{\bar{E},\bar{f}} \|[\ \bar{E} \quad \bar{F}\ ]\|_F, \qquad \text{subject to} \qquad (L_\epsilon + \bar{E}) X_\epsilon = H_\epsilon + \bar{F}, \tag{3.5}$$

then for sufficiently small $\epsilon > 0$, the minimum Frobenius norm solution $X_{t(\epsilon)}$ exists and tends to X_t as ϵ tends to zero.

Proof. To prove the close relation of TLSE solution to the WTLS solution, we need to investigate the right singular vectors of $\widetilde{L}_\epsilon = [L_\epsilon \quad H_\epsilon]$ corresponding to small singular values. Notice that $\widetilde{L}_\epsilon^T$ and $[\widetilde{C}^T \quad \epsilon\widetilde{A}^T]$ have the same left singular vectors, while their singular values are identical up to multiplication by ϵ^{-1}.

To apply Lemma 3.2, let $\widetilde{C}^T = V_C S_C U_C^T$ be the skinny SVD of the full column-rank matrix $\widetilde{C}^T$, and the SVD of $\widetilde{A}\widetilde{Q}_2$ be given by (2.6). It is obvious that $(I_{n+d} - \widetilde{C}^T \widetilde{C}^{\dagger T})\widetilde{A}^T$ has the SVD:

$$(I_{n+d} - \widetilde{C}^T \widetilde{C}^{\dagger T})\widetilde{A}^T = \widetilde{Q}_2 \widetilde{Q}_2^T \widetilde{A}^T = (\widetilde{Q}_2 \widetilde{V})\widetilde{\Sigma}^T \widetilde{U}^T = \overline{V}\widetilde{\Sigma}^T \widetilde{U}^T.$$

By Lemma 3.2, we know that the left and right singular matrices $\widetilde{V}_\epsilon$, $\widetilde{U}_\epsilon$ of $[\widetilde{C}^T \quad \epsilon\widetilde{A}^T]$ satisfies

$$\widetilde{V}_\epsilon = \Big[\underbrace{V_C + \mathcal{O}(\epsilon^2)}_{p} \quad \underbrace{\overline{V} + \mathcal{O}(\epsilon^2)}_{n+d-p}\Big], \qquad \widetilde{U}_\epsilon = \Big[\underbrace{\mathbf{P}_\epsilon\Big(U_C + \mathcal{O}(\epsilon^2)\Big)}_{p} \quad \underbrace{\mathbf{Q}_\epsilon\Big(\widetilde{U} + \mathcal{O}(\epsilon^2)\Big)}_{n+d-p}\Big], \tag{3.6}$$

where $\overline{V} = \widetilde{Q}_2\widetilde{V}_2$,

$$\mathbf{P}_\epsilon = \begin{bmatrix} I_p \\ \epsilon(\widetilde{A}\widetilde{C}^\dagger) \end{bmatrix}, \qquad \mathbf{Q}_\epsilon = \begin{bmatrix} -\epsilon(\widetilde{A}\widetilde{C}^\dagger)^T \\ I_q \end{bmatrix}, \tag{3.7}$$

and the corresponding singular values are just diagonal entries of $S_C + \mathcal{O}(\epsilon^2)$, $\epsilon\widetilde{\Sigma} + \mathcal{O}(\epsilon^3)$. Therefore the SVD of $\widetilde{L}_\epsilon$ is given by $\widetilde{L}_\epsilon = \widetilde{U}_\epsilon\widetilde{S}_\epsilon\widetilde{V}_\epsilon^T$ with

$$\widetilde{S}_\epsilon = \mathrm{diag}(\epsilon^{-1}S_C + \mathcal{O}(\epsilon), \widetilde{\Sigma} + \mathcal{O}(\epsilon^2)), \tag{3.8}$$

and for sufficently small ϵ, the smallest $n+d-p$ singular values of $\widetilde{L}_\epsilon$ can be approximated by $\widetilde{\sigma}_i + \mathcal{O}(\epsilon^2)$ for $i = 1, \ldots, n+d-p$, with a gap between $\tilde{\sigma}_k^2 + \mathcal{O}(\epsilon^2)$ and $\tilde{\sigma}_{k+1}^2 + \mathcal{O}(\epsilon^2)$. Moreover, the bottom right $d \times (n+d-t)$ submatrix in $\widetilde{V}_\epsilon$ has full row rank. Therefore the minimum Frobenius norm WTLS solution $X_{t(\epsilon)}$ to problem (3.5) exists, and in the limit, it takes the form

$$\lim_{\epsilon\to 0+} X_{t(\epsilon)} = \lim_{\epsilon\to 0+} \Big[-(\overline{V}_{12} + \mathcal{O}(\epsilon^2))(\overline{V}_{22} + \mathcal{O}(\epsilon^2))^\dagger \Big] = -\overline{V}_{12}\overline{V}_{22}^\dagger,$$

which is exactly X_t. This leads to the desired results. ■

Remark 3.4. The weighting method of treating TLSE as the WTLS problem (3.4)-(3.5) allows the application of any numerical algorithms for standard TLS model to the WTLS problem, say for a TLSE model arising in color image deblurring and denoising problem (See Example 5.1), we may use the randomized truncated TLS (RTTLS) algorithm [48, Section 3.2.1] on $\widetilde{L}_\epsilon$ to approximate well its large singular values and corresponding singular vectors, from which the RTTLS solution X_{rttls} is approximated by $X_{t(\epsilon)} = \left(\widetilde{V}_{11(\epsilon)}^T\right)^\dagger \widetilde{V}_{21(\epsilon)}^T$ according to the formula (2.3). The idea of computing X_{rttls} is to treat $n+d-t$ small singular values of $\widetilde{L}_\epsilon$ as identical [48], and the parameter $t = k+p$ plays the role of the regularization parameter to guarantee a well-conditioned $\widetilde{V}_{11(\epsilon)}$. A feasible t_* can be chosen based on the L-shaped curve for

$$(t, \quad \log_{10} y_t), \quad \text{with} \quad y_t = \frac{\|H_\epsilon - L_\epsilon X_{t(\epsilon)}\|_F^2}{d + \|X_{t(\epsilon)}\|_F^2}. \tag{3.9}$$

Here the quantity y_t approximates small singular values of $\widetilde{L}_\epsilon$ and the choice t is to avoid small singular value gaps that might lead to a (nearly) singular $\widetilde{V}_{11(\epsilon)}$ (see [41, Theorem 3.14]).

4. Condition numbers of the multidimensional TLSE problem. Condition numbers measure the sensitivity of the solution to the original data in problems, and they play an important role in numerical analysis. In this section, we evaluate the condition number of the multidimensional TLSE problem.

Let $m = p+q$. Define the mapping $\phi : \mathbb{R}^{m(n+d)} \to \mathbb{R}^{nd}$ for the multidimensional TLSE problem (1.2):

$$\phi(c) = \mathrm{vec}(X_t), \quad c = \mathrm{vec}([L \quad H]), \quad \text{for} \quad L = \begin{bmatrix} C \\ A \end{bmatrix}, \qquad H = \begin{bmatrix} D \\ B \end{bmatrix},$$

and the absolute normwise, relative normwise, mixed and componentwise condition numbers of X_t as follows

$$\begin{aligned}
\kappa^{\rm abs}(X_t,L,H) &= \lim_{\epsilon\to 0}\sup\left\{\frac{\|\Delta X_t\|_F}{\|[\Delta L \quad \Delta H]\|_F} : \|[\Delta L \quad \Delta H]\|_F \le \epsilon\|[L \quad H]\|_F\right\},\\
\kappa^{\rm rel}(X_t,L,H) &= \lim_{\epsilon\to 0}\sup\left\{\frac{\|\Delta X_t\|_F}{\epsilon\|X_t\|_F} : \|[\Delta L \quad \Delta H]\|_F \le \epsilon\|[L \quad H]\|_F\right\},\\
m(X_t,L,H) &= \lim_{\epsilon\to 0}\sup\left\{\frac{\|\Delta X_t\|_{\max}}{\epsilon\|X_t\|_{\max}} : |\Delta L| \le \epsilon|L|, \quad |\Delta H| \le \epsilon|H|\right\},\\
c(X_t,L,H) &= \lim_{\epsilon\to 0}\sup\left\{\frac{1}{\epsilon}\left\|\frac{\Delta X_t}{X_t}\right\|_{\max} : |\Delta L| \le \epsilon|L|, \quad |\Delta H| \le \epsilon|H|\right\},
\end{aligned}$$

where $|\cdot|$ denotes the componentwise absolute value, $Y \le Z$ means $y_{ij} \le z_{ij}$ for all i,j, and $\frac{Y}{Z}$ is the entry-wise division defined by $\frac{Y}{Z} := [\frac{y_{ij}}{z_{ij}}]$ and $\frac{\xi}{0}$ is interpreted as zero if $\xi = 0$ and infinity otherwise.

If $\mathrm{vec}(X_t) = \phi(c)$ is continuous and Fréchet differentiable at the neighbourhood of the point c, according to the concept and formulae in [5, 11, 12, 36], the above condition numbers can be formulated as follows:

$$\begin{aligned}
\kappa^{\rm abs}(X_t,L,H) &= \|\phi'(c)\|_2, \qquad \kappa^{\rm rel}(X_t,L,H) = \frac{\|\phi'(c)\|_2\|c\|_2}{\|\phi(c)\|_2},\\
m(X_t,L,H) &= \frac{\||\phi'(c)|\cdot|c|\|_\infty}{\|\phi(c)\|_\infty}, \qquad c(X_t,L,H) = \left\|\frac{|\phi'(c)|\cdot|c|}{|\phi(c)|}\right\|_\infty.
\end{aligned}$$

4.1. Normwise condition number. Notice that $\phi'(c)$ is vital for above condition numbers, while a simple and Fréchet differentiable expression of $\phi(c)$ is not easy to derive. To get $\phi'(c)$, as did in [25], we start from the differentiability of the weighted TLS solution $X_{t(\epsilon)}$ by defining the mapping for the multidimensional WTLS problem (3.4)-(3.5): $\mathrm{vec}(X_{t(\epsilon)}) = \varphi(c_\epsilon)$ for $c_\epsilon = \mathrm{vec}([L_\epsilon \quad H_\epsilon])$. Then we get the first order perturbation estimate $\mathrm{vec}(\Delta X_{t(\epsilon)})$ of WTLS solution based on the result in (2.4), from which the first order perturbation estimate of the TLSE solution is derived by taking the limit $\epsilon \to 0$. Similar limit technique to perform perturbation and condition number analysis of a problem was also used in [27, 34, 46] for equality constrained least squares problem, and in [52] for mixed least squares-total least squares problem.

Theorem 4.1. *With the notations in (2.5)-(2.6), let the skinny SVD of $\widetilde{C}$ be $\widetilde{C} = U_C S_C V_C^T$. Assume that the condition (2.7) holds and the partition $\overline{V}_{22}$ in (2.8) is of full row rank. Denote*

$$\mathbf{P} = \begin{bmatrix} I_p \\ 0_{q\times p}\end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} -(\widetilde{A}\widetilde{C}^\dagger)^T \\ I_q\end{bmatrix}, \quad S_1 = \begin{bmatrix} S_C & 0 \\ 0 & \widetilde{\Sigma}_1\end{bmatrix}, \quad \widehat{V}_1 = \underset{\qquad p \qquad\quad k}{[V_C \quad \overline{V}_1]} = \begin{array}{c} n \\ d \end{array}\underset{t}{\begin{bmatrix} \widehat{V}_{11} \\ \hline \widehat{V}_{21}\end{bmatrix}}.$$

Then for sufficiently small perturbation $\|[\Delta L \quad \Delta H]\|_F$, the first order perturbation estimate for the minimum Frobenius norm TLSE solution $X_t = -\overline{V}_{12}\overline{V}_{22}^\dagger$ takes the form

$$\mathrm{vec}(\Delta X_t) = K\mathrm{vec}([\Delta L \quad \Delta H]) + \mathcal{O}(\|[\Delta L \quad \Delta H]\|_F^2), \tag{4.1}$$

where $K=(H_1+H_2)G\widehat{Z}$ *is exactly the Fréchet derivative* $\phi'(c)$ *and with* $F_{\overline{V}_{22}}=I-\overline{V}_{22}^{\dagger}\overline{V}_{22}$,

$$
\begin{aligned}
&H_1=\Big((\overline{V}_{22}\overline{V}_{22}^T)^{-1}\widehat{V}_{21}\Big)\otimes(\overline{V}_{12}F_{\overline{V}_{22}}), \qquad H_2=\Big(\overline{V}_{22}^{\dagger T}\otimes\widehat{V}_{11}^{\dagger T}\Big)\Pi_{(n+d-t,t)},\\
&G=(S_1^2\otimes I_{n+d-t}-\begin{bmatrix}0_p & 0\\ 0 & I_k\end{bmatrix}\otimes(\widetilde{\Sigma}_2^T\widetilde{\Sigma}_2))^{-1}\big[I_t\otimes\widetilde{\Sigma}_2^T \qquad S_1\otimes I_{n+d-t}\big],\\
&\widehat{Z}=\begin{bmatrix}[0_{(n+d)\times p}\quad \overline{V}_1]^T\otimes(\mathbf{Q}\widetilde{U}_2)^T\\ \Pi_{(t,n+d-t)}\Big(\overline{V}_2^T\otimes[\mathbf{P}U_C\quad \mathbf{Q}\widetilde{U}_1]^T\Big)\end{bmatrix}.
\end{aligned}
\tag{4.2}
$$

Proof. Assume that the SVD of $\widetilde{L}_\epsilon=[L_\epsilon\quad H_\epsilon]=\widetilde{U}_\epsilon\widetilde{S}_\epsilon\widetilde{V}_\epsilon^T$ is given by (3.6)–(3.8), whose factors have partitions as

$$
\begin{aligned}
&\widetilde{U}_\epsilon=\Big[\underset{t}{\widetilde{U}_{1(\epsilon)}}\mid\underset{n+d-t}{\widetilde{U}_{2(\epsilon)}}\Big]=\Big[\underset{p}{\mathbf{P}_\epsilon U_C}\quad \underset{k}{\mathbf{Q}_\epsilon\widetilde{U}_1}\;\Big|\;\underset{n+d-t}{\mathbf{Q}_\epsilon\widetilde{U}_2}\Big]+\mathcal{O}(\epsilon^2),\\
&\widetilde{V}_\epsilon=\begin{array}{c}n\\ d\end{array}\left[\begin{array}{c|c}\widetilde{V}_{11(\epsilon)} & \widetilde{V}_{12(\epsilon)}\\ \widetilde{V}_{21(\epsilon)} & \widetilde{V}_{22(\epsilon)}\end{array}\right]_{\;t\quad n+d-t}=\Big[\underbrace{V_C\quad \overline{V}_1}_{t}\;\Big|\;\underset{n+d-t}{\overline{V}_2}\Big]+\mathcal{O}(\epsilon^2),\\
&\widetilde{S}_{1(\epsilon)}=\mathrm{diag}(\epsilon^{-1}S_C+\mathcal{O}(\epsilon),\ \widetilde{\Sigma}_1+\mathcal{O}(\epsilon^2)),\quad \widetilde{S}_{2(\epsilon)}=\widetilde{\Sigma}_2+\mathcal{O}(\epsilon^2).
\end{aligned}
\tag{4.3}
$$

By applying the result in (2.4) for the WTLS problem (3.4)–(3.5), the first order perturbation estimate of the WTLS solution $X_{t(\epsilon)}$ satisfies

$$
\mathrm{vec}(\Delta X_{t(\epsilon)})\ =(H_{1(\epsilon)}+H_{2(\epsilon)})D_\epsilon Z_\epsilon\mathrm{vec}([\Delta L_\epsilon\quad \Delta H_\epsilon])+\mathcal{O}(\|\Delta L_\epsilon\|_F^2+\|\Delta H_\epsilon\|_F^2),
\tag{4.4}
$$

where with $F_{\widetilde{V}_{22(\epsilon)}}=I-\widetilde{V}_{22(\epsilon)}^{\dagger}\widetilde{V}_{22(\epsilon)}$,

$$
\begin{aligned}
&H_{1(\epsilon)}=\Big((\widetilde{V}_{22(\epsilon)}\widetilde{V}_{22(\epsilon)}^T)^{-1}\widetilde{V}_{21(\epsilon)}\Big)\otimes(\widetilde{V}_{12(\epsilon)}F_{\widetilde{V}_{22(\epsilon)}}),\quad H_{2(\epsilon)}=\Big(\widetilde{V}_{22(\epsilon)}^{\dagger T}\otimes\widetilde{V}_{11(\epsilon)}^{\dagger T}\Big)\Pi_{(n+d-t,t)},\\
&D_\epsilon=(\widetilde{S}_{1(\epsilon)}^2\otimes I_{n+d-t}-I_t\otimes(\widetilde{S}_{2(\epsilon)}^T\widetilde{S}_{2(\epsilon)}))^{-1}\big[I_t\otimes\widetilde{S}_{2(\epsilon)}^T\qquad \widetilde{S}_{1(\epsilon)}\otimes I_{n+d-t}\big],\\
&Z_\epsilon=\begin{bmatrix}\widetilde{V}_{1(\epsilon)}^T\otimes\widetilde{U}_{2(\epsilon)}^T\\ \Pi_{(t,n+d-t)}(\widetilde{V}_{2(\epsilon)}^T\otimes\widetilde{U}_{1(\epsilon)}^T)\end{bmatrix}.
\end{aligned}
\tag{4.5}
$$

In (4.4), $D_\epsilon Z_\epsilon\mathrm{vec}([\Delta L_\epsilon\quad \Delta H_\epsilon])=D_\epsilon Z_\epsilon(I_{n+d}\otimes W_\epsilon^{-1})\mathrm{vec}([\Delta L\quad \Delta H])$. By setting $\widehat{W}_\epsilon=\mathrm{diag}(\epsilon I_p,I_k)$, $\widehat{S}_{1(\epsilon)}=\widehat{W}_\epsilon\widetilde{S}_{1(\epsilon)}$, we obtain $D_\epsilon Z_\epsilon(I_{n+d}\otimes W_\epsilon^{-1})=G_\epsilon\widehat{Z}_\epsilon$ for

$$
\begin{aligned}
&D_\epsilon=(\widehat{S}_{1(\epsilon)}^{\,2}\otimes I_{n+d-t}-\widehat{W}_\epsilon^2\otimes(\widetilde{S}_{2(\epsilon)}^T\widetilde{S}_{2(\epsilon)}))^{-1}\big[\widehat{W}_\epsilon^2\otimes\widetilde{S}_{2(\epsilon)}^T\qquad (\widehat{W}_\epsilon\widehat{S}_{1(\epsilon)})\otimes I_{n+d-t}\big],\\
&G_\epsilon=(\widehat{S}_{1(\epsilon)}^{\,2}\otimes I_{n+d-t}-\widehat{W}_\epsilon^2\otimes(\widetilde{S}_{2(\epsilon)}^T\widetilde{S}_{2(\epsilon)}))^{-1}\big[I_t\otimes\widetilde{S}_{2(\epsilon)}^T\qquad \widehat{S}_{1(\epsilon)}\otimes I_{n+d-t}\big],\\
&\widehat{Z}_\epsilon=\begin{bmatrix}(\widetilde{V}_{1(\epsilon)}\widehat{W}_\epsilon^2)^T\otimes(\widetilde{U}_{2(\epsilon)}^TW_\epsilon^{-1})\\ \Pi_{(t,n+d-t)}\Big(\widetilde{V}_{2(\epsilon)}^T\otimes(W_\epsilon^{-1}\widetilde{U}_{1(\epsilon)}\widehat{W}_\epsilon)^T\Big)\end{bmatrix}.
\end{aligned}
\tag{4.6}
$$

By the expressions in (4.3) and taking the limit $\epsilon\to 0$ for $H_{1(\epsilon)},H_{2(\epsilon)},G_\epsilon$ and $\widehat{Z}_\epsilon$ in (4.5)–(4.6), we obtain the corresponding limit matrices $H_1,H_2,G,\widehat{Z}$ as (4.2), and that $K=(H_1+H_2)G\widehat{Z}$ is exactly the Fréchet derivative $\phi'(c)$. ∎

Theorem 4.2. *With the notations in Theorem 4.1 , the absolute and relative condition numbers of the minimum Frobenius norm TLSE solution X_t are given by*

$$\kappa^{\rm abs}(X_t, L, H) = \|(H_1 + H_2)G\overline{Z}\|_2, \qquad \kappa^{\rm rel}(X_t, L, H) = \|(H_1 + H_2)G\overline{Z}\|_2 \frac{\|[L \quad H]\|_F}{\|X_t\|_F},$$

where

$$H_1 = \big((\overline{V}_{22}\overline{V}_{22}^T)^{-1}\widehat{V}_{21}\big) \otimes (\overline{V}_{12} + X_t\overline{V}_{22}), \quad H_2 = \Big(\overline{V}_{22}^{\dagger T} \otimes (\widehat{V}_{11} + X_t\widehat{V}_{21})\Big)\Pi_{(n+d-t,t)},$$
$$\overline{Z} = \mathrm{diag}\Big(\begin{bmatrix} 0_p & 0 \\ 0 & I_k \end{bmatrix} \otimes (\widetilde{U}_2^T\mathbf{Q}^T) \ , \ \begin{bmatrix} I_p & 0 \\ -\widetilde{U}_1^T(\widetilde{A}\widetilde{C}^\dagger)U_C & I_k \end{bmatrix} \otimes I_{n+d-t} \Big).$$

In particular, when $k = n - p$, H_1 diminishes to zero and $H_2 = \big(\overline{V}_{22}^{-T} \otimes \widehat{V}_{11}^{-T}\big)\Pi_{(d,n)}$ for $\widehat{V}_{11}^{-T} = \widehat{V}_{11} + X_n\widehat{V}_{21}$.

Proof. By the condition number formulae, the absolute and relative condition numbers of the solution X_t are given by

$$\kappa^{\rm abs}(X_t, L, H) = \|\phi'(c)\|_2 = \|K\|_2, \qquad \kappa^{\rm rel}(X_t, L, H) = \frac{\|\phi'(c)\|_2\|c\|_2}{\|X_t\|_F} = \frac{\|K\|_2\|[L \quad H]\|_F}{\|X_t\|_F},$$

in which $\|K\|_2 = \|KK^T\|_2^{1/2} = \|(H_1 + H_2)G\widehat{Z}\widehat{Z}^TG^T(H_1 + H_2)\|^{1/2}$ for

$$\widehat{Z}\widehat{Z}^T = \begin{bmatrix} \begin{bmatrix} 0_p & 0 \\ 0 & I_k \end{bmatrix} \otimes (\widetilde{U}_2^T\mathbf{Q}^T\mathbf{Q}\widetilde{U}_2) & 0 \\ 0 & \Pi_{(t,n+d-t)}\Big(I_{n+d-t} \otimes M\Big)\Pi_{(t,n+d-t)}^T \end{bmatrix} = \breve{Z}\breve{Z}^T,$$

and

$$M = \begin{bmatrix} I_p & U_C^T\mathbf{P}^T\mathbf{Q}\widetilde{U}_1 \\ \widetilde{U}_1^T\mathbf{Q}^T\mathbf{P}U_C & \widetilde{U}_1^T\mathbf{Q}^T\mathbf{Q}\widetilde{U}_1 \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ -\widetilde{U}_1^T(\widetilde{A}\widetilde{C}^\dagger)U_C & I_k \end{bmatrix} \begin{bmatrix} I_p & 0 \\ -\widetilde{U}_1^T(\widetilde{A}\widetilde{C}^\dagger)U_C & I_k \end{bmatrix}^T,$$
$$\breve{Z} = \mathrm{diag}\Big(\begin{bmatrix} 0_p & 0 \\ 0 & I_k \end{bmatrix} \otimes (\widetilde{U}_2^T\mathbf{Q}^T) \ , \ \Pi_{(t,n+d-t)}\Big(I_{n+d-t} \otimes \begin{bmatrix} I_p & 0 \\ -\widetilde{U}_1^T(\widetilde{A}\widetilde{C}^\dagger)U_C & I_k \end{bmatrix}\Big) \Big)$$
$$= \mathrm{diag}\Big(\begin{bmatrix} 0_p & 0 \\ 0 & I_k \end{bmatrix} \otimes (\widetilde{U}_2^T\mathbf{Q}^T) \ , \ \Big(\begin{bmatrix} I_p & 0 \\ -\widetilde{U}_1^T(\widetilde{A}\widetilde{C}^\dagger)U_C & I_k \end{bmatrix} \otimes I_{n+d-t}\Big)\Pi_{(t,n+d-t)} \Big).$$

Notice that $\Pi_{(t,n+d-t)}$ is an orthogonal matrix, then $\|K\|_2 = \|(H_1 + H_2)G\breve{Z}\|_2 = \|(H_1 + H_2)G\overline{Z}\|_2$.

The expression for H_1 is obvious since $X_t = -\overline{V}_{12}\overline{V}_{22}^\dagger$. For H_2, note that the SVDs of $\widetilde{A}\widetilde{Q}_2\widetilde{Q}_2^T = \widetilde{U}\widetilde{\Sigma}\overline{V}^T$ and $\widetilde{C} = U_CS_CV_C^T$ imply $\mathcal{R}(V_C) = \mathcal{R}(\widetilde{C}^T) = \mathcal{R}(\widetilde{Q}_1)$ and $\mathcal{R}(\overline{V}) \subseteq \mathcal{R}(\widetilde{Q}_2)$, therefore $V_C^T\overline{V} = 0$ and

$$\breve{V} := [V_C \quad \overline{V}] = \begin{bmatrix} \widehat{V}_{11} & \overline{V}_{12} \\ \widehat{V}_{21} & \overline{V}_{22} \end{bmatrix} \tag{4.7}$$

is an $(n+d)\times(n+d)$ orthogonal matrix. According to Lemma 2.1(b),

$$\widehat{V}_{11}^{\dagger^T} = \widehat{V}_{11} + X_t \widehat{V}_{21}. \tag{4.8}$$

The formula for H_2 is as desired. Moreover when $k = n-p$, $\overline{V}_{22}$ is nonsingular and so is $\widehat{V}_{11}$ by Lemma 2.1(a). Obviously, $X_n = -\overline{V}_{12}\overline{V}_{22}^{-1}$, and $H_1 = 0$. ■

The simplified form of $\|K\|_2$ still involves Kronecker product operations, which might lead to large storage cost. The following theorem gives compact upper bounds for the normwise condition number.

Theorem 4.3. *Let*

$$\rho_{AC}^{(1)} = 1 + \|\widetilde{C}\|_2 + \|\widetilde{A}\widetilde{C}^\dagger\widetilde{C}\|_2, \quad \rho_{AC}^{(2)} = 1 + \|\widetilde{C}^\dagger\|_2 + \|\widetilde{A}\widetilde{C}^\dagger\|_2, \quad \eta_k^\sigma = \max\left\{1, \frac{\sqrt{\widetilde{\sigma}_k^2+\widetilde{\sigma}_{k+1}^2}}{\widetilde{\sigma}_k^2-\widetilde{\sigma}_{k+1}^2}\right\},$$

then for the absolute normwise condition number, we have

$$\kappa^{\mathrm{abs}}(X_t, L, H) \le (1+\|X_t\|_2^2)\rho_{AC}^{(2)}\eta_k^\sigma.$$

In particular, when $k = n-p$, it has the bounds as

$$\frac{\eta_k^\sigma}{\|\widehat{V}_{11}\|_2\|\overline{V}_{22}\|_2\rho_{AC}^{(1)}} \le \kappa^{\mathrm{abs}}(X_n, L, H) \le (1+\|X_n\|_2^2)\rho_{AC}^{(2)}\eta_k^\sigma.$$

Proof. Let $W_0 = \begin{bmatrix} 0_p & 0 \\ 0 & I_k \end{bmatrix}$. It follows that $\overline{Z} = \Gamma\ddot{Z}$ for $\Gamma = \mathrm{diag}(\Gamma_1,\Gamma_2)\otimes I_{n+d-t}$, with $\Gamma_1 = W_0, \Gamma_2 = \mathrm{diag}(S_C, I_k)$, and

$$\ddot{Z} = \mathrm{diag}\Big(\ W_0\otimes(\widetilde{U}_2^T\mathbf{Q}^T)\ ,\ \begin{bmatrix} S_C^{-1} & 0 \\ -\widetilde{U}_1^T(\widetilde{A}\widetilde{C}^\dagger)U_C & I_k \end{bmatrix}\otimes I_{n+d-t}\ \Big) =: \mathrm{diag}(\ddot{Z}_{11}, \ddot{Z}_{22}).$$

Therefore

$$\kappa^{\mathrm{abs}}(X_t, L, H) \le \|H_1+H_2\|_2\|\overline{G}\|_2\|\ddot{Z}\|_2, \tag{4.9}$$

where $\overline{G} = G\Gamma$, and

$$\overline{G}\,\overline{G}^T = \left(\begin{bmatrix} S_C^4 & 0 \\ 0 & \widetilde{\Sigma}_1^T\widetilde{\Sigma}_1 \end{bmatrix}\otimes I_{n+d-t} + W_0\otimes(\widetilde{\Sigma}_2^T\widetilde{\Sigma}_2)\right)\left(S_1^2\otimes I_{n+d-t} - W_0\otimes(\widetilde{\Sigma}_2^T\widetilde{\Sigma}_2)\right)^{-2}$$

consists of $(k+1)$ diagonal blocks $D^{(i)}$ for $i = 0,1,\cdots,k$ satisfying

$$D^{(0)} = I_{p(n+d-t)}, \qquad D^{(i)} = \mathrm{diag}\Big(\frac{\widetilde{\sigma}_i^2+\widetilde{\sigma}_{k+j}^2}{(\widetilde{\sigma}_i^2-\widetilde{\sigma}_{k+j}^2)^2}\Big), \quad 1\le i\le k, 1\le j\le n+d-t.$$

Note that $\frac{\sigma^2+\eta^2}{(\sigma^2-\eta^2)^2}$ is an increasing function of η and a decreasing function of σ for $\sigma>\eta>0$, therefore

$$\|\overline{G}\|_2 = \|\overline{G}\,\overline{G}^T\|_2^{1/2} = \eta_k^\sigma. \tag{4.10}$$

For the upper bound of $\|H_1 + H_2\|_2$, note that $\check{V}$ in (4.7) is an orthogonal matrix and $X_t = -\overline{V}_{12}\overline{V}_{22}^{\dagger}$. Then by a similar technique in [50, Theorem 3.6], we have

$$\|H_1 + H_2\|_2 \leq \|\overline{V}_{22}^{\dagger}\|_2^2 = 1 + \|X_t\|_2^2. \tag{4.11}$$

For the norm of $\ddot{Z}$ and $\ddot{Z}^{\dagger}$, note that

$$\|\ddot{Z}_{11}\|_2 \leq \|\mathbf{Q}\|_2 = 1 + \|\widetilde{A}\widetilde{C}^{\dagger}\|_2, \quad \|\ddot{Z}_{11}^{\dagger}\|_2 = (\sigma_{\min}(\mathbf{Q}\widetilde{U}_2))^{-1} = (\sigma_{\min}(I + \widetilde{U}_2^T\widehat{C}^T\widehat{C}\widetilde{U}_2))^{-1/2} \leq 1 \tag{4.12}$$

for $\widehat{C} = (\widetilde{A}\widetilde{C}^{\dagger})^T$. Moreover, with $(\widetilde{A}\widetilde{C}^{\dagger})U_C S_C = \widetilde{A}V_C$ and $V_C V_C^T = \widetilde{C}^{\dagger}\widetilde{C}$,

$$\|\ddot{Z}_{22}\|_2 = \left\| \begin{bmatrix} S_C^{-1} & 0 \\ -\widetilde{U}_1^T(\widetilde{A}\widetilde{C}^{\dagger})U_C & I_k \end{bmatrix} \right\|_2 \leq 1 + \|\widetilde{C}^{\dagger}\|_2 + \|\widetilde{A}\widetilde{C}^{\dagger}\|_2 = \rho_{AC}^{(2)},$$
$$\|\ddot{Z}_{22}^{-1}\|_2 = \left\| \begin{bmatrix} S_C V_C^T & 0 \\ \widetilde{U}_1^T\widetilde{A}V_C V_C^T & I_k \end{bmatrix} \right\|_2 \leq 1 + \|\widetilde{C}\|_2 + \|\widetilde{A}\widetilde{C}^{\dagger}\widetilde{C}\|_2 = \rho_{AC}^{(1)}.$$

Therefore $\|\ddot{Z}\|_2 \leq \rho_{AC}^{(2)}$, $\|\ddot{Z}^{\dagger}\|_2 \leq \rho_{AC}^{(1)}$. Combining this with (4.10)–(4.12), the upper bound for $\kappa^{\text{abs}}(X_t, L, H)$ follows.

When $k = n - p$, $H_1 = 0$ and $\kappa^{\text{abs}}(X_n, L, H)$ has the lower bound as

$$\begin{aligned} \kappa^{\text{abs}}(X_n, L, H) &\geq \sigma_{\min}(H_2)\|\overline{G}\ddot{Z}\|_2 \geq \sigma_{\min}(H_2)\sigma_{\min}(\ddot{Z})\|\overline{G}\|_2 \\ &= \frac{\eta_k^{\sigma}}{\|H_2^{-1}\|_2\|\ddot{Z}^{\dagger}\|_2} \geq \frac{\eta_k^{\sigma}}{\|\widehat{V}_{11}\|_2\|\overline{V}_{22}\|_2\rho_{AC}^{(1)}}, \end{aligned}$$

which completes the proof. ■

Remark 4.4. When $\widetilde{C} = 0$ and $\eta_k^{\sigma} > 1$, the upper and lower bounds in Theorem 4.3 reduce to the ones for the TLS problem in [50]. Moreover, note that $\|\overline{V}_{22}^{\dagger}\|_2^2 = 1 + \|X_t\|_2^2 = 1/\sigma_{\min}^2(\overline{V}_{22})$, and it follows from Theorem 4.3 that the multidimensional TLSE problem might be ill-conditioned, when $\widetilde{C}$ is ill conditioned, or the gap between $\widetilde{\sigma}_k$ and $\widetilde{\sigma}_{k+1}$ or $\sigma_{\min}(\overline{V}_{22})$ is small or the solution norm $\|X_t\|_2$ is large.

4.2. Mixed and componentwise condition numbers. For the mixed and componentwise condition numbers, we have the following results.

Theorem 4.5. *With the notations in Theorem 4.1, we have mixed and componentwise condition formulae of X_t as follows:*

$$m(X_t, L, H) = \frac{\||MN|\text{vec}([|L| \quad |H|])\|_{\infty}}{\|X_t\|_{\max}}, \quad c(X_t, L, H) = \left\| \frac{|MN|\text{vec}([|L| \quad |H|])}{\text{vec}(|X_t|)} \right\|_{\infty}, \tag{4.13}$$

where $M = (H_1 + H_2)D^{-1}$, $N = N_1 + N_2$ for

$$D = S_1^2 \otimes I_{n+d-t} - \begin{bmatrix} 0_p & 0 \\ 0 & I_k \end{bmatrix} \otimes (\widetilde{\Sigma}_2^T\widetilde{\Sigma}_2),$$
$$N_1 = [0_{(n+d)\times p} \quad \overline{V}_1]^T \otimes (\mathbf{Q}\widetilde{U}_2\widetilde{\Sigma}_2)^T, \quad N_2 = \Pi_{(t,n+d-t)}\Big(\overline{V}_2^T \otimes [\mathbf{P}U_C S_C \quad \mathbf{Q}\widetilde{U}_1\widetilde{\Sigma}_1]^T\Big).$$

Moreover, if the $(n+d-t)\times t$ *matrix* Y *satisfies* $D\mathrm{vec}(Y)=\mathrm{vec}(\Upsilon)$ *for*

$$\Upsilon = |\mathbf{Q}\widetilde{U}_2\widetilde{\Sigma}_2|^T[|L| \quad |H|][0_{(n+d)\times p} \quad |\overline{V}_1|] + |\overline{V}_2^T|[|L| \quad |H|]^T|[\mathbf{P}U_CS_C \quad \mathbf{Q}\widetilde{U}_1\widetilde{\Sigma}_1]|, \tag{4.14}$$

then the condition numbers have compact upper bounds as

$$m^u(X_t,L,H) = \frac{\left\| |\widehat{V}_{11}^{\dagger T}|Y^T|\overline{V}_{22}^{\dagger}| + |\overline{V}_{12}F_{\overline{V}_{22}}|Y|\widehat{V}_{21}^T(\overline{V}_{22}\overline{V}_{22}^T)^{-1}| \right\|_{\max}}{\|X_t\|_{\max}},$$
$$c^u(X_t,L,H) = \left\| \frac{|\widehat{V}_{11}^{\dagger T}|Y^T|\overline{V}_{22}^{\dagger}| + |\overline{V}_{12}F_{\overline{V}_{22}}|Y|\widehat{V}_{21}^T(\overline{V}_{22}\overline{V}_{22}^T)^{-1}|}{X_t} \right\|_{\max},$$

where the matrix Y *can be formulated via its* i*-th column* $y_i = Ye_i$*:*

$$y_i = (s_i^2 I_{n+d-t} - \tau_i \widetilde{\Sigma}_2^T\widetilde{\Sigma}_2)^{-1}\Upsilon e_i, \quad i=1,2,\cdots,t. \tag{4.15}$$

with s_i *being the* i*-th diagonal element of* S_1 *and* $\tau_i = 1$ *for* $i>p$ *and zero otherwise.*

Proof. By Theorem 4.1 and the concept of condition numbers, the mixed condition number of X_t can be formulated

$$\begin{aligned}
m(X_t,L,H) &= \frac{\||\phi'(c)|\cdot|c|\|_\infty}{\|\phi(c)\|_\infty} = \frac{\left\| |(H_1+H_2)G\widehat{Z}|\cdot \mathrm{vec}([|L| \quad |H|]) \right\|_\infty}{\|X_t\|_{\max}} \\
&= \frac{\left\| \left| M\Big(N_1 + \Pi_{(t,n+d-t)}(I_{n+d-t}\otimes S_1)\big(\overline{V}_2^T \otimes [\mathbf{P}U_C \quad \mathbf{Q}\widetilde{U}_1]^T\big)\right|\Big)\mathrm{vec}([|L| \quad |H|]) \right\|_\infty}{\|X_t\|_{\max}} \\
&= \frac{\||MN|\mathrm{vec}([|L| \quad |H|])\|_\infty}{\|X_t\|_{\max}},
\end{aligned}$$

where the numerator is bounded by

$$\begin{aligned}
&|MN|\mathrm{vec}([|L| \quad |H|]) \le (|H_1|+|H_2|)D^{-1}(|N_1|+|N_2|)\mathrm{vec}([|L| \quad |H|]) \\
&\le (|H_1|+|H_2|)D^{-1}\mathrm{vec}(\Upsilon) \le \mathrm{vec}\Big(|\widehat{V}_{11}^{\dagger T}|Y^T|\overline{V}_{22}^{\dagger}| + |\overline{V}_{12}F_{\overline{V}_{22}}|Y|\widehat{V}_{21}^T(\overline{V}_{22}\overline{V}_{22}^T)^{-1}|\Big),
\end{aligned}$$

where $D\mathrm{vec}(Y) = \mathrm{vec}(\Upsilon)$ gives $YS_1^2 - \widetilde{\Sigma}_2^T\widetilde{\Sigma}_2 Y \begin{bmatrix} 0_p & 0 \\ 0 & I_k \end{bmatrix} = \Upsilon$, leading to the expression of $y_i = Ye_i$ in (4.15).

For the componentwise condition number, it follows that

$$c(X_t,L,H) = \left\| \frac{|\phi'(c)|\cdot|c|}{\phi(c)} \right\|_\infty = \left\| \frac{|MN|\mathrm{vec}([|L| \quad |H|])}{\mathrm{vec}(|X_t|)} \right\|_\infty,$$

and the upper bound $c^u(X_t,L,H)$ follows obviously. ■

The following result is straightforward from Theorem 4.5.

Corollary 4.6. *With the notations in Theorem 4.1 , if for* $k=n-p$, $\widetilde{\sigma}_k > \widetilde{\sigma}_{k+1}$ *and* $\overline{V}_{22}$ *is nonsingular, then the mixed and componentwise condition numbers for the TLSE solution* $X_n = -\overline{V}_{12}\overline{V}_{22}^{-1}$ *satisfy*

$$m(X_n,L,H) = \frac{\||MN|\mathrm{vec}([|L| \quad |H|])\|_\infty}{\|X_n\|_{\max}}, \quad c(X_n,L,H) = \left\| \frac{|MN|\mathrm{vec}([|L| \quad |H|])}{\mathrm{vec}(|X_n|)} \right\|_\infty,$$

where

$$M = (\overline{V}_{22}^{-T} \otimes \widehat{V}_{11}^{-T})\Pi_{(d,n)}(S_1^2 \otimes I_d - \begin{bmatrix} 0_p & 0 \\ 0 & I_{n-p} \end{bmatrix} \otimes (\widetilde{\Sigma}_2^T \widetilde{\Sigma}_2))^{-1}, \tag{4.16}$$

$$N = [0_{(n+d)\times p} \quad \overline{V}_1]^T \otimes (\mathbf{Q}\widetilde{U}_2\widetilde{\Sigma}_2)^T + \Pi_{(n,d)}\Big(\overline{V}_2^T \otimes [\mathbf{P}U_C S_C \quad \mathbf{Q}\widetilde{U}_1\widetilde{\Sigma}_1]^T\Big). \tag{4.17}$$

Moreover, they have upper bounds as

$$m^u(X_n, L, H) = \frac{\Big\| |\overline{V}_{11}^{-T}| Y^T |\overline{V}_{22}^{-1}| \Big\|_{\max}}{\|X_n\|_{\max}}, \qquad c^u(X_n, L, H) = \left\| \frac{|\overline{V}_{11}^{-T}| Y^T |\overline{V}_{22}^{-1}|}{X_n} \right\|_{\max},$$

where the i-th columns of Y and Υ are given by (4.14)-(4.15) with $t = n$.

Theorem 4.7. *For the TLSE problem $Ax \approx \mathsf{b}$ subject to $Cx = \mathsf{d}$, assume that for $k = n - p$, $\widetilde{\sigma}_k > \widetilde{\sigma}_{k+1}$ and $\overline{V}_{22}$ is nonzero. Then for matrices M, N given by Corollary 4.6 and the solution x_n, we have the relation*

$$K = MN = T_1 G(x_n) - T_2,$$

where $G(x) = [x^T \quad -1] \otimes I_{p+q}$,

$$T_1 = 2\rho^{-2}\mathcal{K}x_n u^T - [C_A^\dagger \quad \mathcal{K}A^T], \quad T_2 = \mathcal{K}\Big([I_n \quad 0_{n\times 1}] \otimes u^T\Big),$$

for $C_A^\dagger = (I_n - \mathcal{K}A^T A)C^\dagger$, $\rho = \sqrt{1 + \|x_n\|_2^2}$ and $u^T = \big[-r^T(\widetilde{A}\widetilde{C}^\dagger) \quad r^T\big]$ with $r = Ax - \mathsf{b}$.

Proof. Note that when $k = n - p, d = 1$, in (4.16) and (4.17), $\Pi_{(d,n)} = \Pi_{(n,d)} = I_n$, $\widetilde{\Sigma}_2^T\widetilde{\Sigma}_2 = \widetilde{\sigma}_{n-p+1}^2$ and

$$D^{-1} := (S_1^2 \otimes I_d - \begin{bmatrix} 0_p & 0 \\ 0 & I_{n-p} \end{bmatrix} \otimes (\widetilde{\Sigma}_2^T\widetilde{\Sigma}_2))^{-1} = \mathrm{diag}(S_C^{-2}, (\widetilde{\Sigma}_1^2 - \widetilde{\sigma}_{n-p+1}^2 I_{n-p})^{-1}).$$

Following (4.8), the matrices M, N in (4.16)–(4.17) take the form

$$\begin{aligned} M &= \overline{V}_{22}^{-T}\widehat{V}_{11}^{-T}D^{-1} = \overline{V}_{22}^{-T}[I_n \quad x_n]\widehat{V}_1 D^{-1} = \overline{V}_{22}^{-T}[I_n \quad x_n]\widehat{V}_1 \mathrm{diag}(S_C^{-2}, (\widetilde{\Sigma}_1^2 - \widetilde{\sigma}_{n-p+1}^2 I_{n-p})^{-1}), \\ N &= N_1 + N_2 = [0_{(n+1)\times p} \quad \overline{V}_1]^T \otimes (\mathbf{Q}\widetilde{U}_2\widetilde{\Sigma}_2)^T + \overline{V}_2^T \otimes \Big([\mathbf{P}U_C S_C \quad \mathbf{Q}\widetilde{U}_1\widetilde{\Sigma}_1]^T\Big), \end{aligned}$$

where $\widehat{V}_1 = [\widehat{V}_{11}^T \quad \widehat{V}_{21}^T]^T$ and $\widehat{V}_{11}$, $\widehat{V}_{21}$ are defined in Theorem 4.1 .

In the following, we first derive an equivalent formula for $(\widetilde{\Sigma}_1^2 - \widetilde{\sigma}_{n-p+1}^2 I_{n-p})^{-1}$. Let $\widetilde{Q}_2$ be given by (2.14), based on the SVD in (2.6): $\widetilde{A}\widetilde{Q}_2 = \widetilde{U}_1\widetilde{\Sigma}_1\widetilde{V}_1^T + \widetilde{U}_2\widetilde{\Sigma}_2\widetilde{V}_2^T$, partition

$$\widetilde{V} = [\widetilde{V}_1 \quad \widetilde{V}_2] \quad = \quad \begin{array}{c c} & \begin{array}{cc} n-p & 1 \end{array} \\ \begin{array}{c} n-p \\ 1 \end{array} & \begin{bmatrix} \widetilde{V}_{11} & \widetilde{V}_{12} \\ \widetilde{V}_{21} & \widetilde{V}_{22} \end{bmatrix} \end{array}.$$

Note that AQ_2 is the first $n-p$ columns of $\widetilde{A}\widetilde{Q}_2$, therefore $AQ_2 = \widetilde{U}_1\widetilde{\Sigma}_1\widetilde{V}_{11}^T + \widetilde{U}_2\widetilde{\Sigma}_2\widetilde{V}_{12}^T$, from which

$$\begin{aligned} Y_0 &:= (AQ_2)^T(AQ_2) - \widetilde{\sigma}_{n-p+1}^2 I_{n-p} \\ &= [\widetilde{V}_{11} \quad \widetilde{V}_{12}]\mathrm{diag}(\widetilde{\Sigma}_1^2 - \widetilde{\sigma}_{n-p+1}^2 I_{n-p}, 0)[\widetilde{V}_{11} \quad \widetilde{V}_{12}]^T = \widetilde{V}_{11}(\widetilde{\Sigma}_1^2 - \widetilde{\sigma}_{n-p+1}^2 I_{n-p})\widetilde{V}_{11}^T, \end{aligned} \tag{4.18}$$

in which $\widetilde{V}_{1j}$ satisfies

$$\overline{V}_{1j} = Q_2\widetilde{V}_{1j} + \beta^{-1}x_C\widetilde{V}_{2j} = Q_2\widetilde{V}_{1j} - x_C\overline{V}_{2j}, \quad j=1,2, \tag{4.19}$$

according to the relation $\overline{V} = \widetilde{Q}_2\widetilde{V}$. Moreover $\widetilde{V}_{22} = -\beta\overline{V}_{22}$ is nonzero. By Lemma 2.1(a), $\widetilde{V}_{11}$ is nonsingular. From (4.18), we obtain

$$(\Sigma_1^2 - \widetilde{\sigma}_{n-p+1}^2 I_{n-p})^{-1} = \widetilde{V}_{11}^T Y_0^{-1}\widetilde{V}_{11}. \tag{4.20}$$

It should be noted that for any column vector z and matrices M_i,

$$M_1(M_2\otimes z^T) = (M_1M_2)\otimes z^T, \qquad z^T\otimes M_3 = M_3(z^T\otimes I).$$

With $\widehat{V}_1 = [V_C \quad \overline{V}_1]$ and the expressions in (4.20), we therefore obtain

$$\begin{aligned}
MN_1 &= \left([I_n \quad x_n]\widehat{V}_1\begin{bmatrix} 0_{p\times(n+1)} \\ \widetilde{V}_{11}^T Y_0^{-1}\widetilde{V}_{11}\overline{V}_1^T\end{bmatrix}\right)\otimes(\mathbf{Q}\widetilde{U}_2\widetilde{\Sigma}_2\overline{V}_{22}^{-1})^T \\
&= \left([I_n \quad x_n]\overline{V}_1\widetilde{V}_{11}^T Y_0^{-1}\widetilde{V}_{11}\overline{V}_1^T\right)\otimes(\mathbf{Q}\widetilde{U}_2\widetilde{\Sigma}_2\overline{V}_{22}^{-1})^T, \\
MN_2 &= \widehat{V}_{11}^{-T}D^{-1}\Big([-x_n^T \quad 1]\otimes[\mathbf{P}U_CS_C \quad \mathbf{Q}\widetilde{U}_1\widetilde{\Sigma}_1]^T\Big) \\
&= [I_n \quad x_n]\widehat{V}_1D^{-1}[\mathbf{P}U_CS_C \quad \mathbf{Q}\widetilde{U}_1\widetilde{\Sigma}_1]^T\big([-x_n^T \quad 1]\otimes I_{p+q}\big) \\
&= [I_n \quad x_n](\widetilde{C}^\dagger\mathbf{P}^T + \overline{V}_1\widetilde{V}_{11}^T Y_0^{-1}(\widetilde{U}_1\widetilde{\Sigma}_1\widetilde{V}_{11}^T)^T\mathbf{Q}^T)([-x_n^T \quad 1]\otimes I_{p+q}),
\end{aligned}$$

in which

$$\begin{aligned}
\widetilde{U}_2\widetilde{\Sigma}_2\overline{V}_{22}^{-1} &= \widetilde{A}\widetilde{Q}_2\widetilde{V}_2\overline{V}_{22}^{-1} = \widetilde{A}\overline{V}_2\overline{V}_{22}^{-1} = [A \quad \mathsf{b}]\begin{bmatrix}-x_n \\ 1\end{bmatrix} = -r. \\
\widetilde{U}_1\widetilde{\Sigma}_1\widetilde{V}_{11}^T &= \widetilde{U}_1\widetilde{\Sigma}_1\widetilde{V}_1^T\begin{bmatrix}I_{n-p} \\ 0\end{bmatrix} = [(\widetilde{A}\widetilde{Q}_2) - \widetilde{U}_2\widetilde{\Sigma}_2\widetilde{V}_2^T]\begin{bmatrix}I_{n-p} \\ 0\end{bmatrix} \\
&= AQ_2 - \widetilde{U}_2\widetilde{\Sigma}_2\overline{V}_{22}^{-1}\overline{V}_{22}\widetilde{V}_{12}^T = AQ_2 + r\overline{V}_{22}\widetilde{V}_{12}^T,
\end{aligned} \tag{4.21}$$

and

$$\begin{aligned}
\widetilde{V}_{11}\overline{V}_1^T &= [\widetilde{V}_{11}\widetilde{V}_{11}^T \quad \widetilde{V}_{11}\widetilde{V}_{21}^T]\widetilde{Q}_2^T = [I_{n-p} - \widetilde{V}_{12}\widetilde{V}_{12}^T \quad -\widetilde{V}_{12}\widetilde{V}_{22}^T]\widetilde{Q}_2^T \\
&= \Big([I_{n-p} \quad 0] - \widetilde{V}_{12}\widetilde{V}_2^T\Big)\widetilde{Q}_2^T = [Q_2^T \quad 0] - \widetilde{V}_{12}\overline{V}_2^T \\
&= [Q_2^T \quad 0] - \widetilde{V}_{12}\overline{V}_{22}^T[\overline{V}_{22}^{-T}\overline{V}_{12}^T \quad 1] = [Q_2^T \quad 0] + \widetilde{V}_{12}\overline{V}_{22}^T[x_n^T \quad -1].
\end{aligned} \tag{4.22}$$

Therefore

$$[I_n \quad x_n]\overline{V}_1\widetilde{V}_{11}^T = Q_2. \tag{4.23}$$

Moreover, note that the Greville's method [2, Chapter 7, Section 5] gives

$$\widetilde{C}^\dagger = \begin{bmatrix}\left(I_n - \omega^{-1}x_Cx_C^T\right)C^\dagger \\ \omega^{-1}x_C^TC^\dagger\end{bmatrix}, \qquad \omega = 1 + \|x_C\|_2^2.$$

Combining this with the expression for x_n in (2.16) with $\mathcal{K} = Q_2 Y_0^{-1} Q_2^T$, $r_C = Ax_C - \mathsf{b}$, we have

$$
\begin{aligned}
&[I_n \quad x_n]\widetilde{C}^\dagger - \mathcal{K}A^T\widetilde{A}\widetilde{C}^\dagger \\
&= (I_n - \omega^{-1}\mathcal{K}A^T r_C x_C^T)C^\dagger - \mathcal{K}A^T(A - \omega^{-1} r_C x_C^T)C^\dagger = (I_n - \mathcal{K}A^T A)C^\dagger.
\end{aligned} \tag{4.24}
$$

Besides, according to (4.19), the matrix $\mathcal{K}$ applying on $Q_2\widetilde{V}_{12}\overline{V}_{22}^T$ gives

$$
\mathcal{K}Q_2\widetilde{V}_{12}\overline{V}_{22}^T = \mathcal{K}(\overline{V}_{12} + x_C\overline{V}_{22})\overline{V}_{22}^T = \mathcal{K}(-x_n + x_C)\overline{V}_{22}\overline{V}_{22}^T = -\rho^{-2}\mathcal{K}x_n, \tag{4.25}
$$

where $\overline{V}_{22}\overline{V}_{22}^T = \|\overline{V}_{22}\|_2^2 = \rho^{-2}$ based on the fact that $[x_n^T \quad -1] = \rho\big[\overline{V}_{12}^T \quad \overline{V}_{22}^T\big]^T$ for $\rho^2 = 1 + \|x_n\|_2^2$, and $\mathcal{K}(x_C - x_n) = -\mathcal{K}x_n$ since $Q_2^T x_C = 0$.

Combining (4.21)-(4.25) with $u^T = r^T\mathbf{Q}^T = \big[-r^T(\widetilde{A}\widetilde{C}^\dagger) \quad r^T\big]$, we have

$$
\begin{aligned}
MN_1 &= -\Big[Q_2 Y_0^{-1} Q_2^T\Big([I_n \quad 0] + Q_2\widetilde{V}_{12}\overline{V}_{22}^T[x_n^T \quad -1]\Big)\Big] \otimes u^T \\
&= -\Big(\mathcal{K}[I_n \quad 0] - \rho^{-2}\mathcal{K}x_n[x_n^T \quad -1]\Big) \otimes u^T \\
&= -\mathcal{K}([I_n \quad 0] \otimes u^T) + (\rho^{-2}\mathcal{K}x_n u^T)([x_n^T \quad -1]) \otimes I_{p+q}), \\
MN_2 &= \Big[[I_n \quad x_n]\widetilde{C}^\dagger \quad 0\Big] + \mathcal{K}A^T[-\widetilde{A}\widetilde{C}^\dagger \quad I_q] + \mathcal{K}Q_2\widetilde{V}_{12}\overline{V}_{22}^T u^T\Big)([-x_n^T \quad 1] \otimes I_{p+q}) \\
&= \Big([(I_n - \mathcal{K}A^T A)C^\dagger \quad \mathcal{K}A^T] - \rho^{-2}\mathcal{K}x_n u^T\Big)\big([-x_n^T \quad 1] \otimes I_{p+q}\big).
\end{aligned}
$$

Consequently,

$$
M(N_1 + N_2) = \Big(2\rho^{-2}\mathcal{K}x_n u^T - [(I_n - \mathcal{K}A^T A)C^\dagger \quad \mathcal{K}A^T]\Big)\big([x_n^T \quad -1] \otimes I_{p+q}\big) - \mathcal{K}\big([I_n \quad 0] \otimes u^T\big),
$$

which is exactly K. The assertion in the theorem then follows. ■

Remark 4.8. In Theorem 4.7 the matrix $K = T_1 G(x_n) - T_2$ characterizing the first order perturbation of the TLSE solution is just the one derived in [25] for the single right-hand-side TLSE problem. As revealed in [24], the single right-hand-side TLSE can be viewed as a generalization of the linear least squares problem with equality constraint (LSE), and the normwise condition numbers of TLSE in [24] includes the ones for LSE problems. Therefore the normwise condition numbers of multidimensional TLSE problem unify the counterparts for the single right-hand-side TLSE [24] and LSE [7] problems.

Remark 4.9. It is observed that the formulae for three types of condition numbers involve the Kronecker product which might lead to large storage and computational cost. For mixed and componentwise condition numbers, we can use their upper bounds as alternatives, while for the normwise condition numbers, as did in [50], we can compute

$$
\kappa^{\mathrm{abs}}(X_t, L, H) = \|\breve{H}\|_2 = \|\breve{H}^T\breve{H}\|_2^{1/2}, \quad \textit{for} \quad \breve{H} = (H_1 + H_2)G\overline{Z} = (H_1 + H_2)D^{-1}\breve{Z},
$$

by applying the power method to the matrix $\breve{H}^T\breve{H}$, in which D is defined in Theorem 4.5, $\breve{Z} = \big[I_t \otimes \widetilde{\Sigma}_2^T \quad S_1 \otimes I_{n+d-t}\big]\overline{Z}$. In the power scheme, the matrix-vector multiplications associated with $\breve{H}$ and $\breve{H}^T$ can be transformed into Kronecker product-free operations, say for $\breve{H}f$, where $f = [f_1^T \quad f_2^T]^T$ with $f_i = \mathrm{vec}(F_i)$ with $F_1 \in \mathbb{R}^{(p+q)\times t}, F_2 \in \mathbb{R}^{(n+d-t)\times t}$,

$$\begin{aligned} g &:= \breve{H}f = (H_1+H_2)D^{-1}\mathrm{vec}\left((\mathbf{Q}\widetilde{U}_2\widetilde{\Sigma}_2)^T F_1 \begin{bmatrix} 0_p & 0 \\ 0 & I_k \end{bmatrix} + F_2 \begin{bmatrix} S_C & -U_C^T(\widetilde{A}\widetilde{C}^\dagger)^T \widetilde{U}_1\widetilde{\Sigma}_1 \\ 0 & \widetilde{\Sigma}_1 \end{bmatrix}\right) \\ &= (H_1+H_2)\mathrm{vec}(T) = \mathrm{vec}\Big((\overline{V}_{12} + X_t\overline{V}_{22}))T\widehat{V}_{21}^T(\overline{V}_{22}\overline{V}_{22}^T)^{-1} + (\overline{V}_{11} + X_t\overline{V}_{21})T^T\overline{V}_{22}^\dagger\Big), \end{aligned}$$

where $t_i = Te_i$ satisfies

$$t_i = (s_i^2 I_{n+d-t} - \tau_i \widetilde{\Sigma}_2^T\widetilde{\Sigma}_2)^{-1}\left((\mathbf{Q}\widetilde{U}_2\widetilde{\Sigma}_2)^T F_1 \begin{bmatrix} 0_p & 0 \\ 0 & I_k \end{bmatrix} + F_2 \begin{bmatrix} S_C & -U_C^T(\widetilde{A}\widetilde{C}^\dagger)^T \widetilde{U}_1\widetilde{\Sigma}_1 \\ 0 & \widetilde{\Sigma}_1 \end{bmatrix}\right) e_i,$$

in which s_i, τ_i are the same as those in Theorem 4.5. The Kronecker product-free expression associated with $\breve{H}^T g$ can be derived in a similar manner. Here we omit these.

Remark 4.10. If the matrix A has a linear structure so that it can be represented by $A = \sum_{i=1}^{\ell} \alpha_i S_i$ and

$$\mathrm{vec}(A) = \sum_{i=1}^{\ell} \alpha_i \mathrm{vec}(S_i) = \Phi_A^{\mathrm{struct}}\boldsymbol{a}, \quad \Phi_A^{\mathrm{struct}} = [\mathrm{vec}(S_1), \mathrm{vec}(S_2), \cdots, \mathrm{vec}(S_\ell)],$$

where $S_1, S_2, \cdots, S_\ell$ are linearly independent basis of structured matrices, and $\boldsymbol{a} = [\alpha_1, \alpha_2, \cdots, \alpha_\ell]^T$. By the statement in [23, Theorem 4.1], Φ_A^{struct} is column orthogonal and has full column rank, with at most one nonzero entry in each row. If B, C, D also have linear structures, then there exist column orthogonal matrices $\Phi_{A,B}^{\mathrm{struct}}, \Phi_{C,D}^{\mathrm{struct}}, \Phi_{L,H}^{\mathrm{struct}}$ and a permutation matrix Π so that

$$\begin{aligned} &\mathrm{vec}([A \quad B]) = \Phi_{A,B}^{\mathrm{struct}} \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix} := \begin{bmatrix} \Phi_A^{\mathrm{struct}} & 0 \\ 0 & \Phi_B^{\mathrm{struct}} \end{bmatrix} \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix}, \\ &\mathrm{vec}([C \quad D]) = \Phi_{C,D}^{\mathrm{struct}} \begin{bmatrix} \boldsymbol{c} \\ \boldsymbol{d} \end{bmatrix}, \quad \mathrm{vec}([L \quad H]) = \Phi_{L,H}^{\mathrm{struct}}[\boldsymbol{c}^T \ \boldsymbol{d}^T \ \boldsymbol{a}^T \ \boldsymbol{b}^T]^T, \end{aligned}$$

where $\Phi_{L,H}^{\mathrm{struct}} = \Pi \mathrm{diag}(\Phi_{C,D}^{\mathrm{struct}}, \Phi_{A,B}^{\mathrm{struct}})$, and Π is the permutation matrix such that $\mathrm{vec}([L \quad H]) = \Pi \begin{bmatrix} \mathrm{vec}([C \ D]) \\ \mathrm{vec}([A \ B]) \end{bmatrix}$.

For the perturbed TLSE problem, if we restrict the perturbation matrices $[\Delta L \quad \Delta H]$ to have the same structure as that of $[L \quad H]$, that is, $\mathrm{vec}([\Delta L \quad \Delta H]) = \Phi_{L,H}^{\mathrm{struct}}\boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon} = [\Delta\boldsymbol{c}^T \ \Delta\boldsymbol{d}^T \ \Delta\boldsymbol{a}^T \ \Delta\boldsymbol{b}^T]^T$. By defining the mapping ϕ such that $\phi(\boldsymbol{c}, \boldsymbol{d}, \boldsymbol{a}, \boldsymbol{b}) = x_{\mathrm{tlse}}$, the first order perturbation result becomes $\Delta x = K\Phi_{L,H}^{\mathrm{struct}}\boldsymbol{\epsilon} + \mathcal{O}(\|\boldsymbol{\epsilon}\|_2^2)$ based on (4.1). According to the concept of condition numbers, the relative normwise, mixed and componentwise condition numbers for structured TLSE take the following forms

$$\kappa^{\text{struct}}(\boldsymbol{c},\boldsymbol{d},\boldsymbol{a},\boldsymbol{b}) = \|K\Phi_{L,H}^{\text{struct}}\|_2 \frac{\|[\boldsymbol{c}^T \ \boldsymbol{d}^T \ \boldsymbol{a}^T \ \boldsymbol{b}^T]\|_2}{\|x_{\text{tlse}}\|_2},$$

$$m^{\text{struct}}(\boldsymbol{c},\boldsymbol{d},\boldsymbol{a},\boldsymbol{b}) = \frac{\left\||K\Phi_{L,H}^{\text{struct}}| \cdot [|\boldsymbol{c}|^T \ |\boldsymbol{d}|^T \ |\boldsymbol{a}|^T \ |\boldsymbol{b}|^T]^T\right\|_\infty}{\|x_{\text{tlse}}\|_\infty},$$

$$c^{\text{struct}}(\boldsymbol{c},\boldsymbol{d},\boldsymbol{a},\boldsymbol{b}) = \left\|\frac{|K\Phi_{L,H}^{\text{struct}}| \cdot [|\boldsymbol{c}|^T \ |\boldsymbol{d}|^T \ |\boldsymbol{a}|^T \ |\boldsymbol{b}|^T]^T}{|x_{\text{tlse}}|}\right\|_\infty.$$

5. Numerical experiments.

In this section, we present numerical examples to verify our results. The following numerical tests are performed via MATLAB with machine precision $u = 2.22e-16$ in a laptop with Intel Core (TM) i5-5200U CPU.

Example 5.1. In this example, we apply the multidimensional TLSE model to solve a color image deblurring problem with constraints on color channels. Based on the RGB color space, a color image is represented by a three-dimensional vector function,

$$\mathbf{u}(x,y) = \begin{bmatrix} u_r(x,y) & u_g(x,y) & u_b(x,y) \end{bmatrix}^T,$$

where $u_r(x,y)$, $u_g(x,y)$ and $u_b(x,y)$ denote the red, green and blue channels, respectively, and (x,y) belongs to a square domain $\Omega \subseteq \mathbb{R}^2$. Let the domain Ω be divided into d non-intersect square sub-domains of same size: $\Omega_1, \cdots, \Omega_d$, that is $\Omega = \cup_{i=1}^d \Omega_i$ and $\Omega_i \cap \Omega_j = \varnothing$ if $i \neq j$. Now we present a new parallel and coupling degradation model of color image blurring and noising:

$$\hat{\mathbf{u}}(x,y) = K \star \mathbf{u}(x,y) + \mathbf{b}(x,y), \quad (x,y) \in \Omega_i, \tag{5.1}$$

where

$$K = \begin{bmatrix} K_{rr} & K_{gr} & K_{br} \\ K_{rg} & K_{gg} & K_{bg} \\ K_{rb} & K_{gb} & K_{bb} \end{bmatrix} \tag{5.2}$$

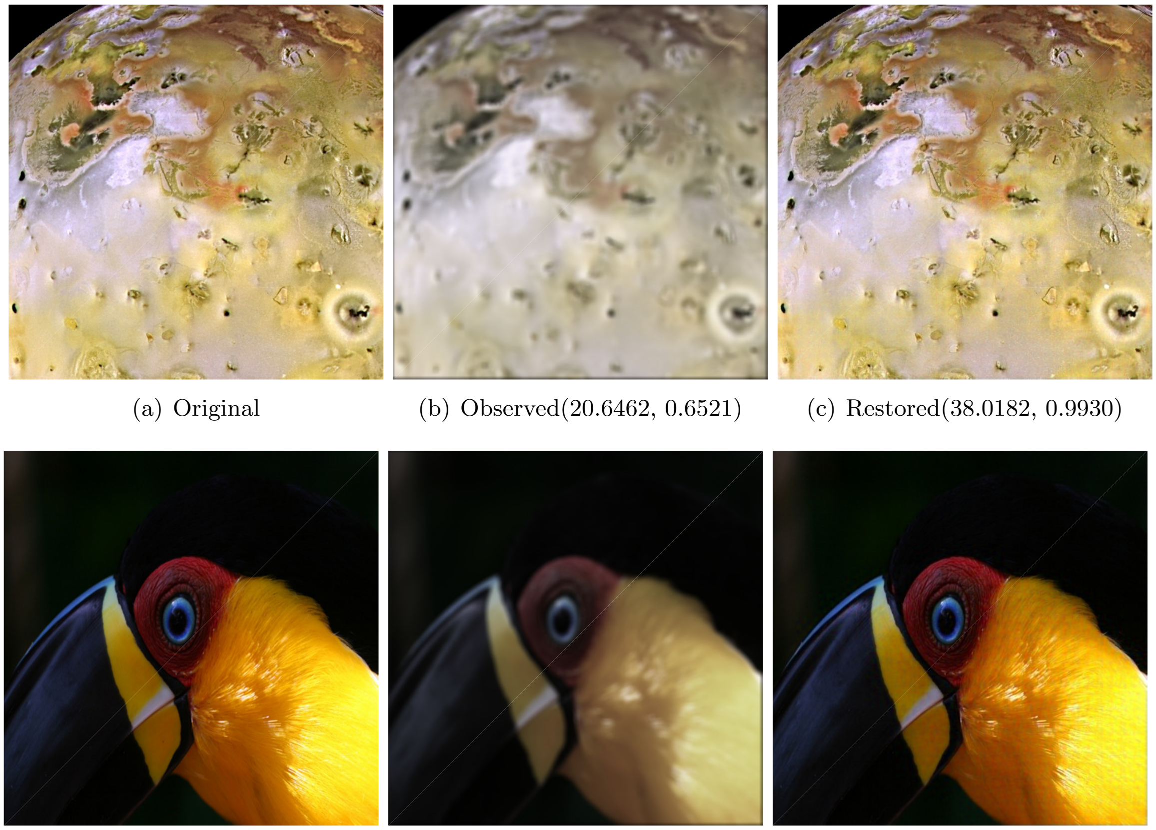
is a coupling blur operator independent of i with each sub-matrices being real blur operators, $\star$ is the convolution operation, and $\mathbf{b}(x,y)$ is an additional Gaussian noise. The left hand side of (5.1), $\hat{\mathbf{u}}(x,y)$, is called an observed color image. The aim of color image deblurring and denoising is to recover the original color image $\mathbf{u}(x,y)$ from such observed color image under the parallel and coupling degradation model (5.1).

Let $A \in \mathbb{R}^{3n\times 3n}$ be the discrete operator of K, and u_i, $b_i \in \mathbb{R}^n$ denote the discrete forms of $\mathbf{u}(x,y)$ and $\mathbf{b}(x,y)$ on Ω_i, respectively. Then a new model of color image deblurring and denoising is proposed by

$$\min_{E\in\mathbb{R}^{3n\times 3n},\ F\in\mathbb{R}^{3n\times d}} \|[E \quad F]\|_F, \quad \text{subject to} \quad (A+E)U = B+F, \quad CU = D, \tag{5.3}$$

where $U = [u_1 \ \cdots \ u_d]$, $B = [b_1 \ \cdots \ b_d] \in \mathbb{R}^{3n\times d}$,

$$C = \begin{bmatrix} I_n & I_n & I_n \end{bmatrix} \in \mathbb{R}^{n\times 3n}$$



(a) Original (b) Observed(20.6462, 0.6521) (c) Restored(38.0182, 0.9930)

(d) Original (e) Observed(19.4510, 0.8113) (f) Restored(35.2479, 0.9835)

Figure 5.1. *The visual comparison of color image prediction: The original color images are in the first column, the observed color images are in the second column, and the restored color images are in the last column. The PSNR and SSIM values of observed and restored images are listed in the brackets under the images.*

and $D \in \mathbb{R}^{n\times d}$ is a binary matrix taking 1 or 0 as its entries.

The Gaussian blur operator A and the Gaussian noise are respectively generated by the Matlab commands: `imfilter` and `randn`. The original color images are shown in the first column of Figure 5.1 and they are of size 512×512, i.e, the square domain is $\Omega = \{(i,j) : 1 \leq i,j \leq 512\}$. For the first image in Figure 5.1, the number of sub-domains is $d = 2^{10}$ and the size of each sub-domain is $n = 2^4 \times 2^4 = 2^8$. In the practical implementation, we add extra boundaries about 8 pixels to each sub-domain and thus extend n to 1024. For the second image in Figure 5.1, the number of sub-domains is $d = 2^8$ and the size of each sub-domain is $n = 2^5 \times 2^5 = 2^{10}$. Similarly, we add extra boundaries about 8 pixels to each sub-domain and thus extend n to 2304.

The deblurring problem is solved via randomized truncated TLS (RTTLS) algorithm [48] applied on the WTLS problem (3.4), in which the weighting factor $\epsilon = 10^{-8}$, and the sample size $l = t + 10$. In Figure 5.2, the L-shaped curves (3.9) are plotted for two observed images, with 10 as the stepsize of t. The regularization parameter t_* is chosen to be the abscissa of the

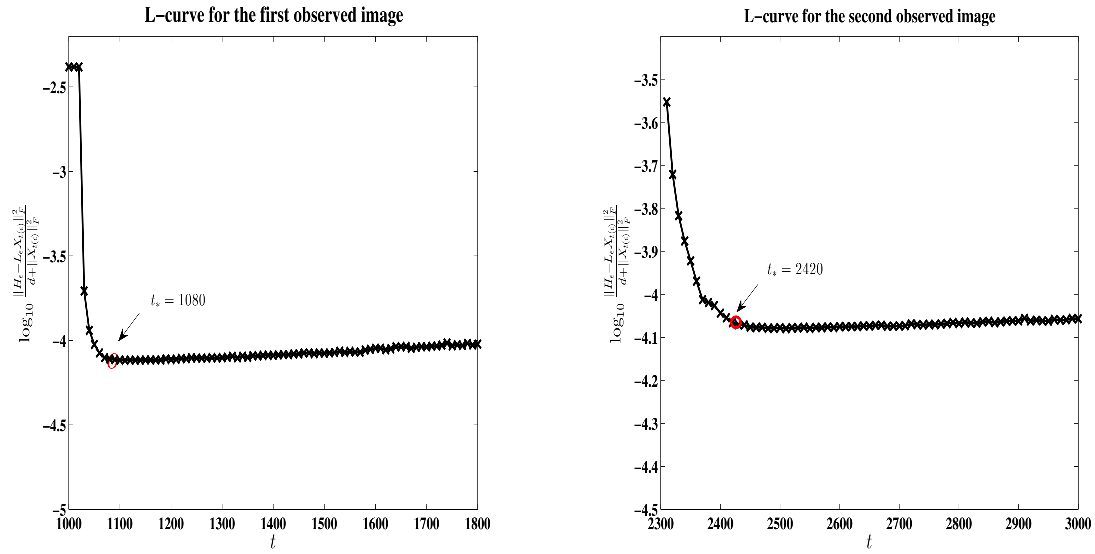


Figure 5.2. *The L-curves of* $\left(t, \log_{10}\frac{\|H_\epsilon - L_\epsilon X_{t(\epsilon)}\|_F^2}{d+\|X_{t(\epsilon)}\|_F^2}\right)$ *for solving TLSE based on the weighting method.*

point near the corner of the L-curve, with the adjacent decay rate of y_t in magnitude smaller than a threshold τ.

By taking $\tau = 10^{-6}$, we obtain $t_* = 1080, 2420$ for restoring two blurred images respectively. The restorations are shown in the third column of Figure 5.1 and their PSNR and SSIM values are listed below. It is observed that our method successfully completes the deblurring and denoising processing and restores color images with high PSNR and SSIM values.

Example 5.2. In this example, we generate small random multidimensional TLSE problems to verify the rationality of the first order perturbation estimate in Theorem 4.1. The entries in $[C \quad D]$ and $[A \quad B]$ are generated as random variables uniformly distributed in the interval (0,1), via Matlab command '$\texttt{rand}(\cdot)$'. Set $p = 10, q = 40, n = 40, d = 5$, and let the perturbations to the data be given by

$$[\Delta C \quad \Delta D] = \epsilon * \texttt{rand}(p, n+d), \qquad [\Delta A \quad \Delta B] = \epsilon * \texttt{rand}(q, n+d).$$

Choose $t = 10, 20, 30, 40$ and compute the solutions to the original and perturbed problems via the QR-SVD method. In Table 5.1, with respect to different ϵ, we compute the absolute error

$$\eta_{\Delta X_t} = \|\mathrm{vec}(\Delta X_t) - K\mathrm{vec}([\Delta L \quad \Delta H])\|_\infty .$$

The tabulated results show that $\eta_{\Delta X_t} = \mathcal{O}(\epsilon^2)$, illustrating the rationality of the first order perturbation estimates in Theorem 4.1.

Example 5.3. In this example, we do some numerical experiments for TLSE from piecewise-polynomial data fitting problem that is modified from [3, Chapter 16] and [7, Example 3], in order to compare the sharpness of three types of condition numbers in evaluating the forward error of the solution.

Table 5.1
The absolute error of the first order perturbation estimate of vec(ΔX_t)

t	10	20	30	40
$\epsilon = 10^{-2}$	1.9e-4	6.3e-4	5.2e-4	3.0e-4
$\epsilon = 10^{-4}$	5.6e-8	2.3e-8	3.7e-8	2.1e-8
$\epsilon = 10^{-6}$	2.7e-12	3.7e-12	1.8e-12	1.5e-12

Given N points (t_i, y_i) on the plane, we are seeking to find a piecewise-polynomial function $f(t)$ fitting the above set of the points, where

$$f(t) = \begin{cases} f_1(t), & t \le a, \\ f_2(t), & t > a, \end{cases}$$

with a given, and $f_1(t)$ and $f_2(t)$ polynomials of degree three or less,

$$f_1(t) = x_1 + x_2 t + x_3 t^2 + x_4 t^3, \qquad f_2(t) = x_5 + x_6 t + x_7 t^2 + x_8 t^3, .$$

The conditions that $f_1(a) = f_2(a)$ and $f_1'(a) = f_2'(a)$ are imposed, so that $f(t)$ is continuous and has a continuous first derivative at $t = a$. Suppose the N data are numbered so that $t_1, \ldots, t_M \le a$ and $t_{M+1}, \ldots, t_N > a$. The conditions $f_1(a) - f_2(a) = 0$ and $f_1'(a) - f_2'(a) = 0$ lead to the equality constraint $Cx = \mathsf{d}$ for $x = [x_1, x_2, \ldots, x_8]^T$ and

$$C = \begin{bmatrix} 1 & a & a^2 & a^3 & -1 & -a & -a^2 & -a^3 \\ 0 & 1 & 2a & 3a^2 & 0 & -1 & -2a & -3a^2 \end{bmatrix}, \quad \mathsf{d} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

The vector x that minimizes the sum of squares of the prediction errors

$$\sum_{i=1}^{M} (f_1(t_i) - y_i)^2 + \sum_{i=M+1}^{N} (f_2(t_i) - y_i)^2,$$

gives $\min_x \|Ax - b\|_2$, where

$$A = \begin{bmatrix} 1 & t_1 & t_1^2 & t_1^3 & 0 & 0 & 0 & 0 \\ 1 & t_2 & t_2^2 & t_2^3 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & t_M & t_M^2 & t_M^3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & t_{M+1} & t_{M+1}^2 & t_{M+1}^3 \\ 0 & 0 & 0 & 0 & 1 & t_{M+2} & t_{M+2}^2 & t_{M+2}^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 & t_N & t_N^2 & t_N^3 \end{bmatrix}, \qquad b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \\ y_{M+1} \\ \vdots \\ y_N \end{bmatrix},$$

and the matrix A is of 50% sparsity. If more than one observation vector is allowed, the data fitting problem becomes the multidimensional TLSE problem (1.2).

Take $M = 200, N = 400$ and sample $t_i \in [0,1]$ randomly. For a randomly generated piecewise-polynomial function $f(t)$ with a predetermined a, we compute the corresponding function value $y_i = f(t_i)$. Since the matrices A, C do not have linear structures, we consider random componentwise perturbations on the data as

$$\Delta L = 10^{-12} \cdot E_{N+2,8} \odot L, \qquad \Delta H = 10^{-12} \cdot E_{N+2,d} \odot H, \tag{5.4}$$

where $E_{s,t}$ is the random $s \times t$ matrix whose entries are uniformly distributed on the interval (0,1), $\odot$ denotes the entrywise multiplication.

For simplicity let $\kappa_{\rm n}, m, c$ denote the relative normwise, mixed and componentwise condition numbers given in Theorem 4.2 and Theorem 4.5, respectively. Set

$$x = \mathrm{vec}(X_t), \quad \epsilon_{\rm n} = \frac{\|[\Delta L \quad \Delta H]\|_F}{\|[L \quad H]\|_F}, \qquad \epsilon_{\rm c} = \min\{\epsilon : |\Delta L| \le \epsilon |L|, |\Delta H| \le \epsilon |H|\},$$

where t is a random integer between p and n such that $\overline{V}_{22}$ is of full row rank, and the quantity $\epsilon_{\rm n}$ is used to evaluate the upper bound of the forward error $\frac{\|\Delta x\|_2}{\|x\|_2}$ via $\epsilon_{\rm n}\kappa_{\rm n}$, while $\epsilon_{\rm c}$ is to derive upper bounds for $\frac{\|\Delta x\|_\infty}{\|x\|_\infty}, \|\frac{\Delta x}{x}\|_\infty$ via mixed and componentwise condition numbers. Moreover we let $\rho = \rho_{AC}^{(2)} \eta_k^\sigma$ be the factor for upper bounds of $\kappa^{\rm abs}(X_t, L, H)$.

We list numerical results with respect to different a, and for each a, we generate two different problems and compare the estimated upper bound with actual relative forward errors. It is observed that for fixed a, the problems with a larger $\|X_t\|_2^2$ and moderate ρ produce larger condition number estimates, which illustrates that the norm $\|X_t\|_2$ is a factor to affect the condition number of TLSE problem. However, whether $\|X_t\|_2$ is big or small, the estimated upper bounds of the forward error via $\epsilon_{\rm n}\kappa_{\rm n}$, $\epsilon_{\rm c} m$, $\epsilon_{\rm c} c$ are about one or two orders of magnitude larger than the actual forward error of the solution. Among three upper bounds $\kappa_{\rm n}^u, m^u, c^u$ of condition numbers, the normwise condition number-based upper bound $\kappa_{\rm n}^u$ is acceptable and is about one or two orders of magnitude larger than $\kappa_{\rm n}$. The upper bounds m^u, c^u are sharper, which are at most one order of magnitude larger than the corresponding exact condition numbers, therefore they are good estimates of corresponding condition numbers.

Example 5.4. This example is modified from [1]. We test how the ill-conditioning of $\widetilde{C}$ and small singular values gap affect the condition numbers and the accuracy of the solutions. Set $p = d = 5, n = 10, q = 20, k = 3, t = p + k = 8$, and let $\widetilde{Q}$ be an arbitrary $(n+d) \times (n+d)$ orthogonal matrix and $\widetilde{Q}_1$ be the submatrix of $\widetilde{Q}$ by taking its first p columns. Let U_0 be an arbitrary $p \times p$ orthogonal matrix, y, z be unit column vectors of length $q, n+d$, respectively. Set

$$\begin{aligned}
&\widetilde{C} = [C \quad D] = U_0 \mathrm{diag}([1, 0.5, 0.1, 0.1, \kappa_C^{-1}])\widetilde{Q}_1^T, \quad \widetilde{A} = [A \ B] = \hat{A}\widetilde{Q}^T, \quad \text{with} \\
&\hat{A} = (I_q - 2yy^T)[\hat{\Sigma} \quad O](I_{n+d} - 2zz^T), \\
&\hat{\Sigma} = \mathrm{diag}(10, 8, 1, 1, 1, 1, 1, 1 - \delta/2, 1 - \delta, 1 - 2\delta, 1/6, 1/7, \cdots, 1/10]),
\end{aligned}$$

where κ_C is used to control the condition number of $[C \ D]$. Note that $\widetilde{A}\widetilde{Q}_2$ is the last $n+d-p$ columns of $\widetilde{A}\widetilde{Q}$, and by the interlacing theorem of the singular values, the relation $1 = \sigma_j(\widetilde{A}\widetilde{Q}) \ge \sigma_j(\widetilde{A}\widetilde{Q}_2) \ge \sigma_{p+j}(\widetilde{A}\widetilde{Q})$, for $j = k, k+1$ and therefore $0 < \delta < 5/12$ can be used to control the gap of the singular values $\widetilde{\sigma}_k, \widetilde{\sigma}_{k+1}$ of $\widetilde{A}\widetilde{Q}_2$.

Table 5.2

Comparisons of forward errors and upper bounds for the perturbed TLSE problem

a	$\|X_t\|_2^2$	ρ	$\frac{\|\Delta x\|_2}{\|x\|_2}$	$\epsilon_{\rm n}\kappa_{\rm n}$	$\epsilon_{\rm n}\kappa_{\rm n}^u$	$\frac{\|\Delta x\|_\infty}{\|x\|_\infty}$	$\epsilon_{\rm c}m$	$\epsilon_{\rm c}m^u$	$\|\frac{\Delta x}{x}\|_\infty$	$\epsilon_{\rm c}c$	$\epsilon_{\rm c}c^u$
0.1	4.2	12.0	2.2e-13	2.1e-11	7.7e-10	2.6e-13	3.8e-12	1.2e-11	7.3e-13	6.9e-12	3.2e-11
	1.6e5	76.0	1.1e-11	3.0e-9	7.7e-7	9.8e-12	9.1e-10	1.7e-9	1.5e-10	2.1e-8	2.8e-8
0.3	4.2	12.0	1.2e-13	2.1e-11	7.5e-10	1.7e-13	4.0e-12	1.4e-11	5.2e-13	7.2e-12	3.1e-11
	2.4e5	42.0	7.4e-12	3.5e-9	5.2e-7	6.4e-12	1.2e-9	3.0e-9	4.3e-11	6.4e-9	1.3e-8
0.5	5.6	12.0	1.8e-13	4.7e-11	8.2e-10	2.3e-13	1.2e-11	3.3e-11	2.7e-11	2.0e-9	7.2e-9
	5.3e4	68.0	5.8e-12	1.0e-9	3.8e-7	6.8e-12	7.8e-10	1.6e-9	1.2e-9	1.0e-7	1.6e-7
0.7	3.0	11.0	1.2e-13	2.2e-11	7.1e-10	1.4e-13	4.8e-12	1.6e-11	2.7e-13	8.4e-12	3.0e-11
	1.3e7	75.0	1.4e-10	1.5e-8	7.0e-6	1.4e-10	6.5e-9	1.1e-8	2.3e-8	1.1e-6	1.9e-6
0.9	2.3	11.0	4.0e-14	2.2e-11	6.9e-10	5.8e-14	5.7e-12	1.9e-11	8.0e-14	7.9e-12	2.8e-11
	5.2e8	42.0	1.0e-9	1.0e-7	2.7e-5	1.3e-9	7.6e-8	2.1e-7	1.7e-9	5.4e-7	1.0e-6

Consider the same perturbation as in (5.4), for different κ_C and δ, we compute the forward errors and upper bounds via three types condition numbers in Table 5.3. It is observed that the estimated upper bounds of the forward errors via $\epsilon_{\rm n}\kappa_{\rm n}$, $\epsilon_{\rm c}m$, $\epsilon_{\rm c}c$ are about one or two orders of magnitude larger than the corresponding forward errors of the solutions, even the quantity ρ is very large. For the compact upper bounds m^u, c^u of condition numbers, m^u, c^u are very sharp in most cases, while $\kappa_{\rm n}^u$ is not robust against the ill-conditioning of $\widetilde{C}$ and sometimes they are three orders of magnitude larger than $\kappa_{\rm n}$ and five or six orders of magnitude larger than $\frac{\|\Delta x\|_2}{\|x\|_2}$.

Table 5.3

Comparisons of forward error and upper bounds for the perturbed TLSE problem

σ	$\|X_t\|_2^2$	ρ	$\frac{\|\Delta x\|_2}{\|x\|_2}$	$\epsilon_{\rm n}\kappa_{\rm n}$	$\epsilon_{\rm n}\kappa_{\rm n}^u$	$\frac{\|\Delta x\|_\infty}{\|x\|_\infty}$	$\epsilon_{\rm c}m$	$\epsilon_{\rm c}m^u$	$\|\frac{\Delta x}{x}\|_\infty$	$\epsilon_{\rm c}c$	$\epsilon_{\rm c}c^u$
$\kappa_C = 10^1$											
0.1	2.1	2.4e3	2.1e-12	5.4e-10	3.9e-8	1.7e-12	3.4e-11	7.8e-11	3.2e-11	8.6e-10	2.1e-9
0.01	0.71	6.7e3	4.2e-12	2.0e-10	6.5e-8	4.9e-12	1.4e-10	1.5e-10	1.1e-8	1.9e-7	3.0e-7
0.001	0.9	1.6e5	1.1e-10	1.2e-8	2.4e-6	1.4e-10	3.1e-9	4.2e-9	3.6e-10	1.1e-8	1.3e-8
$\kappa_C = 10^3$											
0.1	0.74	1.2e5	1.1e-10	1.5e-8	1.4e-6	1.2e-10	1.1e-9	2.3e-9	6.5e-10	1.4e-8	4.5e-8
0.01	2.1	9.8e5	7.7e-11	8.2e-9	8.7e-6	6.7e-11	6.3e-10	1.4e-9	4.7e-9	3.6e-8	9.5e-8
0.001	0.51	5.5e6	1.2e-10	2.7e-8	7.7e-5	1.5e-10	3.3e-9	5.3e-9	3.6e-9	9.9e-8	2.4e-7
$\kappa_C = 10^6$											
0.1	2.4	8.9e7	2.7e-8	9.7e-6	1.4e-3	3.4e-8	6.2e-7	1.2e-6	6.6e-7	2.7e-5	1.3e-4
0.01	4.8	6.1e8	1.2e-7	3.5e-5	1.2e-2	1.1e-7	1.2e-6	2.9e-6	2.2e-6	2.8e-5	6.7e-5
0.001	2.0	4.6e9	2.0e-8	1.1e-5	6.5e-2	2.1e-8	2.5e-7	4.6e-7	4.6e-7	5.5e-6	1.8e-5

6. Conclusion. In this paper, we investigate the solution of multidimensional TLSE problem and prove that it is equivalent to the multidimensional weighted TLS solution in the limit sense, with the aid of perturbation theory of invariant subspace. Based on this close relation, various numerical algorithms for TLS can be developed for solving TLSE via the weighting method. Moreover, the closed formula for the first order perturbation estimate of the minimum Frobenius norm TLSE solution $X_t = -\overline{V}_{12}\overline{V}_{22}^{\dagger}$ is derived. The expressions for normwise, mixed and componentwise condition numbers of problem TLSE are also presented, as well as their computable upper bounds. All expressions and upper bounds of these condition numbers generalize those for the TLSE problem [25] and multidimensional TLS problem [30].

The effectiveness of the weighting method is shown to solve a color image deblurring and denoising problem. Some numerical examples are also given to demonstrate the effectiveness in estimating the forward errors. Tightness of upper bounds for mixed and componentwise condition numbers are shown in numerical examples, even for ill-conditioned problems, while it is not necessarily tight for the upper bounds of the normwise condition number. Therefore in order to derive good estimates of forward errors via normwise condition number, we recommend using power scheme to compute the true value to avoid Kronecker product operations.

Acknowledgments. The authors are grateful to the handling editor and two anonymous referees for their useful comments and suggestions, which greatly improved the original presentation.

REFERENCES

[1] M. Baboulin And S. Gratton, *A contribution to the conditioning of the total least-squares problem*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 685-699.

[2] A. Ben-Israel And T. N.E. Greville, *Generalized Inverses, Theory and Applications*, 2nd ed., Springer-Verlag, New York, 2003.

[3] S. Boyd And L. Vandenberghe, *Introduction to Applied Linear Algebra-Vectors, Matrices, and Least Squares*, Cambridge University Press, Cambridge, 2018.

[4] A.J. Cox And N.J. Higham, *Accuracy and stability of the null space method for solving the equality constrained least squares problem*, BIT, 39 (1999), pp. 34-50.

[5] F. Cucker, H. Diao And Y. Wei, *On mixed and componentwise condition numbers for Moore-Penrose inverse and linear least squares problems*, Math. Comput., 76 (2007), pp. 947-963.

[6] X. Chang And D. Titley-Peloquin, *Backward perturbation analysis for scaled total least-squares problems*, Numer. Linear Algebra Appl., 16 (2009), pp. 627-648.

[7] H. Diao, *Condition numbers for a linear function of the solution of the linear least squares problem with equality constraints*, J. Comput. Appl. Math., 344 (2018), pp. 640-656.

[8] H. Diao And Y. Sun, *Mixed and componentwise condition numbers for a linear function of the solution of the total least squares problem*, Linear Algebra Appl., 544 (2018), pp. 1-29.

[9] H. Diao, Y. Wei And P. Xie, *Small sample statistical condition estimation for the total least squares problem*, Numer. Algorithms, 75 (2017), pp. 1-21.

[10] E.M. Dowling, R.D. Degroat And D.A. Linebarger, *Total least squares with linear constraints*, IEEE International Conference on Acoustics, 5 (1992), pp. 341-344.

[11] A.J. Geurts, *A contribution to the theory of condition*, Numer. Math., 39 (1982), pp. 85-96.

[12] I. Gohberg And I. Koltracht, *Mixed, componentwise, and structured condition numbers*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 688-704.

[13] S. Gratton, D. Titley-Peloquin And J. T. Ilunga, *Sensitivity and conditioning of the truncated total least squares solution*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1257-1276.

[14] G.H. Golub And C.F. Van Loan, *An analysis of the total least squares problem*, SIAM J. Numer. Anal. Appl., 17 (1980), pp. 883-893.

[15] G.H. Golub And C.F. Van Loan, *Matrix Computations (4ed.)*, Johns Hopkins University Press, Baltimore, 2013.

[16] A. Graham, *Kronecker Products and Matrix Calculus with Application*, Ellis Horwood and John Wiley & Sons, New York, 1981.

[17] K. Hermus, W. Verhelst, P. Lemmerling, P. Wambacq And S. Van Huffel, *Perceptual audio modeling with exponentially damped sinusoids*, Signal Processing, 85 (2005), pp. 163-176.

[18] I. Hnětynková, M. Plešinger And D. Maria Sima, *Solvability of the core problem with multiple right-hand sides in the TLS sense*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 861-876.

[19] Z. Jia And B. Li, *On the condition number of the total least squares problem*, Numer. Math., 125 (2013), pp. 61-87.

[20] A.N. Langville And W.J. Stewart, *The Kronecker product and stochastic automata networks*, J. Comput. Appl. Math., 167 (2004), pp. 429-447.

[21] P. Lemmerling And B. De Moor, *Misfit versus latency*, Automatica, 37 (2001), pp. 2057-2067.

[22] P. Lemmerling, N. Mastronardi And S. Van Huffel, *Efficient implementation of a structured total least squares based speech compression method*, Linear Algebra Appl., 366 (2003), pp. 295-315.

[23] B. Li And Z. Jia, *Some results on condition numbers of the scaled total least squares problem*, Linear Algebra Appl., 435 (2011), pp. 674-686.

[24] Q. Liu, C. Chen And Q. Zhang, *Perturbation analysis for total least squares problems with linear equality constraint*, Appl. Numer. Math., 161 (2021), pp. 69-81.

[25] Q. Liu And Z. Jia, *On condition numbers of the total least squares problem with linear equality constraint*, Numer Algorithms, (2021), https://doi.org/10.1007/s11075-021-01191-w.

[26] Q. Liu, S. Jin, L. Yao And D. Shen, *The revisited total least squares problems with linear equality constraint*, Appl. Numer. Math., 152 (2020), pp. 275-284.

[27] Q. Liu And M. Wei, *On direct elimination methods for solving the equality constrained least squares problem*, Linear Multilinear Algebra, 58 (2010), pp. 173-184.

[28] Q. Meng, H. Diao And Z. Bai, *Condition numbers for the truncated total least squares problem and their estimations*, Numer. Linear Algebra Appl., DOI: 10.1002/nla.2369, 2021.

[29] I. Markovsky And S. Van Huffel, *Overview of total least squares methods*, Signal Processing, 87 (2007), pp. 2283-2302.

[30] L. Meng, B. Zheng and Y. Wei, *Condition numbers of the multi-dimensional total least squares problems having more than one solution*, Numer. Algorithms, 84 (2020), pp. 887-908.

[31] M. Ng, N. Bose And J. Koo, *Constrained total least squares for color image reconstruction*, Total Least Squares and Errors-in-Variables Modelling III: Analysis, Algorithms and Applications, Kluer Academic Publishers, (2002), pp. 365-374.

[32] M. Ng, R. Plemmons And F. Pimentel, *A new approach to constrained total least squares image restoration*, Linear Algebra Appl., 316 (2000), pp. 237-258.

[33] K. Pearson, *On lines and planes of closest fit to systems of points in space*, Phil. Mag., 2 (1901), pp. 559-572.

[34] A. R. De Pieero And M. Wei, *Some new properties of the equality constrained and weighted least squares problem*, Linear Algebra Appl., 320 (2000), pp. 145-165.

[35] M.A. Rahman And K. Yu, *Total least squares approach for frequency estimation using linear prediction*, IEEE Transactions on Acoustics, Speech, and Signal Processing 35 (1987), pp.1440-1454.

[36] J.R. Rice, *A theory of condition*, SIAM J. Numer. Anal., 3 (1966), pp. 287-310.

[37] B. Schaffrin, *A note on constrained total least squares estimation*, Linear Algebra Appl., 417 (2006), pp. 245-258.

[38] G.W. Stewart, *On the asymptotic behavior of scaled singular value and QR decompositions*, Math. Comp., 43 (1984), pp. 483-489.

[39] G. W. Stewart And J.G. Sun, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.

[40] S. Van Huffel, *On the significance of nongeneric total least squares problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 20-35.

[41] S. Van Huffel And J. Vandewalle, *The Total Least Squares Problems: Computational Aspects and Analysis*, Vol. 9 of Frontiers in Applied Mathematics, SIAM, Philadelphia, 1991.

[42] S. Van Huffel And J. Vandewalle, *Analysis and solution of the nongeneric total least squares problem*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 360-372.

[43] M. Wei, *The analysis for the total least squares problem with more than one solution*, SIAM J. Matrix Anal.

Appl., 13 (1992), pp. 746-763.

[44] M. Wei, *Algebraic relations between the total least squares and least squares problems with more than one solution*, Numer. Math., 62 (1992), pp. 123-148.

[45] M. Wei, *Perturbation theory for the rank-deficient equality constrained least squares problem*, SIAM J. Numer. Anal., 29 (1992), pp. 1462-1481.

[46] M. Wei And A. R. De Pieero, *Upper perturbation bounds of weighted projections, weighted and constrained least squares problems*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 931-951.

[47] P. Xie, H. Xiang And Y. Wei, *A contribution to perturbation analysis for total least squares problems*, Numer. Algorithms, 75 (2017), pp. 381-395.

[48] P. Xie, H. Xiang And Y. Wei, *Randomized algorithms for total least squares problems*, Numer. Linear Algebra Appl., 26 (2019), e2219.

[49] L. Zhou, L. Lin, Y. Wei And S. Qiao, *Perturbation analysis and condition numbers of scaled total least squares problems*, Numer. Algorithms, 51 (2009), pp. 381-399.

[50] B. Zheng, L. Meng And Y. Wei, *Condition numbers of the multidimensional total least squares problem*, SIAM J. Matrix Anal. Appl., 38 (2017), pp. 924-948.

[51] L. Zhang And Y. Wei, *Randomized core reduction for discrete ill-posed problem*, J. Comput. Appl. Math., 375 (2020), 112797

[52] B. Zheng And Z. Yang, *Perturbation analysis for mixed least squares-total least squares problems*, Numer. Linear Algebra Appl., 26 (2019), e2239.