

# Partial condition number for the equality constrained linear least squares problem

Hanyu Li<sup>1</sup> · Shaoxin Wang<sup>2</sup>

Received: 30 August 2016 / Accepted: 31 March 2017 / Published online: 7 April 2017  
© Springer-Verlag Italia 2017

**Abstract** In this paper, the normwise condition number of a linear function of the equality constrained linear least squares solution called the partial condition number is considered. Its expression and closed formulae are first presented when the data space and the solution space are measured by the weighted Frobenius norm and the Euclidean norm, respectively. Then, we investigate the corresponding structured partial condition number when the problem is structured. To estimate these condition numbers with high reliability, the probabilistic spectral norm estimator and the small-sample statistical condition estimation method are applied and two algorithms are devised. The obtained results are illustrated by numerical examples.

**Keywords** Linear least squares problem · Equality constraint · Partial condition number · Probabilistic spectral norm estimator · Small-sample statistical condition estimation

**Mathematics Subject Classification** 65F20 · 65F35 · 65F30 · 15A12 · 15A60

---

The work is supported by the National Natural Science Foundation of China (No. 11671060) and the Fundamental Research Funds for the Central Universities (No. 106112015CDJXY100003).

---

✉ Hanyu Li  
lihy.hy@gmail.com; hyl@cq.edu.cn

Shaoxin Wang  
shaoxin.w@gmail.com

<sup>1</sup> College of Mathematics and Statistics, Chongqing University, Chongqing 401331, People's Republic of China

<sup>2</sup> School of Statistics, Qufu Normal University, Qufu 273165, People's Republic of China

### 1 Introduction and preliminaries

The equality constrained linear least squares problem can be stated as follows:

$$\text{LSE : } \min_{Bx=d} \|b - Ax\|_2, \tag{1.1}$$

where  $A \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{s \times n}$  with  $m + s \geq n \geq s \geq 0$ ,  $b \in \mathbb{R}^m$  and  $d \in \mathbb{R}^s$ . Hereafter, the symbols  $\mathbb{R}^{m \times n}$  and  $\mathbb{R}^n$  stand for the set of  $m \times n$  real matrices and the real vector space of dimension  $n$ , respectively. To ensure that the LSE problem (1.1) has a unique solution, we need to assume that [6]

$$\text{rank}(B) = s, \quad \text{null}(A) \cap \text{null}(B) = \{0\}. \tag{1.2}$$

The first condition in (1.2) implies that the linear system  $Bx = d$  is consistent and hence that the LSE problem (1.1) has a solution, and vice versa; The second one, which says that the matrix  $[A^T, B^T]^T$  is full column rank, guarantees that there is a unique solution to (1.1), and vice versa. Here, for a matrix  $C$ ,  $C^T$  denotes its transpose. Throughout this paper, we assume that the conditions in (1.2) always hold. In this case, the unique solution to the LSE problem (1.1) can be written as [6,8]

$$x(A, B, b, d) = (AP)^\dagger b + B_A^\dagger d, \tag{1.3}$$

where

$$P = I_n - B^\dagger B, \quad B_A^\dagger = (I_n - (AP)^\dagger A)B^\dagger$$

with  $I_n$  being the identity matrix of order  $n$  and  $B^\dagger$  being the Moore-Penrose inverse of  $B$ . When  $s = 0$ , i.e.,  $B = 0$  and  $d = 0$ , the LSE problem (1.1) reduces the classic linear least squares (LLS) problem

$$\text{LLS : } \min_{x \in \mathbb{R}^n} \|b - Ax\|_2, \tag{1.4}$$

the conditions in (1.2) reduce to  $A$  being full column rank which ensures that the solution to (1.4) is unique, and the solution (1.3) reduces to  $x(A, b) = A^\dagger b$ .

The LSE problem finds many important applications in some areas. For example, we will encounter it in the analysis of large scale structures, in signal processing, and in solving inequality constrained least squares problem [4,6,19]. So, some scholars considered its algorithms and perturbation analysis (see e.g., [4,6,8,10,19,29]). An upper bound for the normwise condition number of the LSE problem was presented in [8], and the mixed and componentwise condition numbers and their easily computable upper bounds of this problem can be derived from [21] as the special case.

In this paper, we mainly consider the partial condition number of the LSE problem when the data space  $\mathbb{R}^{m \times n} \times \mathbb{R}^{s \times n} \times \mathbb{R}^m \times \mathbb{R}^s$  and the solution space  $\mathbb{R}^n$  are measured by the weighted Frobenius norm

$$\|(\alpha_A A, \alpha_B B, \alpha_b b, \alpha_d d)\|_F = \sqrt{\alpha_A^2 \|A\|_F^2 + \alpha_B^2 \|B\|_F^2 + \alpha_b^2 \|b\|_2^2 + \alpha_d^2 \|d\|_2^2} \tag{1.5}$$

with  $\alpha_A > 0, \alpha_B > 0, \alpha_b > 0,$  and  $\alpha_d > 0,$  and the Euclidean norm  $\|x\|_2,$  respectively. As mentioned in Abstract, the partial condition number which is also called the subspace condition number [7] is referred to the condition number of a linear function of the LSE solution  $x(A, B, b, d),$  i.e.,  $L^T x(A, B, b, d)$  with  $L \in \mathbb{R}^{n \times k}$  ( $k \leq n$ ). This kind of condition number has some advantages. For example, when  $L$  is the identity matrix or a column vector of the identity matrix, the partial condition number will reduce to the condition number of the solution  $x(A, B, b, d)$  or of an element of the solution. Cao and Petzold first proposed the partial condition number for linear systems [7]. Later, it was proposed for LLS problem and total least squares problem [1, 2]. In [1, 2, 7], the authors also provided some specific motivations for investigating this kind of condition number.

The idea on the weighted Frobenius norm can be traced back to Gratton [14], who derived the normwise condition number for the LLS problem (1.4) based on the following weighted Frobenius norm

$$\|(\alpha A, \beta b)\|_F = \sqrt{\alpha^2 \|A\|_F^2 + \beta^2 \|b\|_2^2}, \quad \alpha > 0, \beta > 0. \tag{1.6}$$

Subsequently, this kind of norm was used for the partial condition number for the LLS problem [1] and the normwise condition number of the truncated singular value solution of a linear ill-posed problem [5]. As pointed out in [14], this norm is very flexible. With it, we can monitor the perturbations on  $A$  and  $b$ . For example, if  $\alpha \rightarrow \infty,$  no perturbation on  $A$  will be permitted; similarly, if  $\beta \rightarrow \infty,$  there will be no perturbation on  $b$  allowed. Obviously, the norm in (1.5) is a simple generalization of the one in (1.6), and is also very flexible. There is another kind of generalization of the norm in (1.6):  $\|(TA, \beta b)\|_F,$  which was used by Wei et al. in [30] for the normwise condition number of rank deficient LLS problem. Here,  $T$  is a positive diagonal matrix.

Like the structured linear systems and the structured LLS problem, the structured LSE problem arises in many applications, e.g., in signal processing and the area of optimization [6, 19]. Rump [25, 26] presented the perturbation theory for the structured linear systems with respect to normwise distances and componentwise distances. The obtained results generalized the corresponding ones in [15]. For the structured LLS problems, Xu et al. [31] considered their structured normwise condition numbers, and Cucker and Diao [9] presented their structured mixed and componentwise condition numbers. In addition, the structured condition numbers for the total least squares problem were provided by Li and Jia in [20]. The results in [20, 25, 26, 31] show that the structured condition number can be much tighter than the unstructured one. So, based on the study on the partial condition number, we also investigate the structured partial condition number of the structured LSE problem.

The rest of this paper is organized as follows. Section 2 presents the expression and closed formulae of the partial condition number for the LSE problem. The expression of the corresponding structured partial condition number is given in Sect. 3. On basis of the probabilistic spectral norm estimator by Hochstenbach [16] and the small-sample statistical condition estimation (SSCE) method by Kenney and Laub [18], Sect. 4 is devoted to the statistical estimates and algorithms of the results derived in Sects. 2

and 3. The numerical experiments for illustrating the obtained results are provided in Sect. 5.

Before moving to the following sections, we first introduce some results on the operator ‘vec’ and Kronecker product, and the generalized singular value decomposition (GSVD) of a matrix pair. They will be necessary later in this paper.

For a matrix  $A = [a_1, \dots, a_n] \in \mathbb{R}^{m \times n}$  with  $a_i \in \mathbb{R}^m$ , the operator ‘vec’ is defined as follows

$$\text{vec}(A) = [a_1^T, \dots, a_n^T]^T \in \mathbb{R}^{mn}.$$

Let  $A = (a_{ij}) \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{p \times q}$ . The *Kronecker product* between  $A$  and  $B$  is defined by (see, e.g., [17, Chapter 4])

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \in \mathbb{R}^{mp \times nq}.$$

This definition implies that when  $m = 1$  and  $q = 1$ , i.e., when  $A$  is a row vector and  $B$  is a column vector,

$$A \otimes B = BA. \tag{1.7}$$

From [17, Chapter 4], we have

$$(A \otimes B)^T = (A^T \otimes B^T), \tag{1.8}$$

$$\text{vec}(AXB) = (B^T \otimes A) \text{vec}(X), \tag{1.9}$$

$$\begin{aligned} \Pi_{mn} \text{vec}(A) &= \text{vec}(A^T), \\ \Pi_{pm}(A \otimes B) \Pi_{nq} &= (B \otimes A), \end{aligned} \tag{1.10}$$

where  $X \in \mathbb{R}^{n \times p}$ , and  $\Pi_{mn}$  is the *vec-permutation matrix* depending only on the orders  $m$  and  $n$ . Especially, when  $n = 1$ , i.e.,  $A$  is a column vector, then  $\Pi_{nq} = I_q$  and hence

$$\Pi_{pm}(A \otimes B) = (B \otimes A). \tag{1.11}$$

In addition, the following result is also from [17, Chapter 4]

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD), \tag{1.12}$$

where the matrices  $C$  and  $D$  are of suitable orders.

For the matrix pair  $A, B$  in (1.1) and (1.2), there exist orthogonal matrices  $U \in \mathbb{R}^{m \times m}$  and  $V \in \mathbb{R}^{s \times s}$ , and a nonsingular matrix  $X \in \mathbb{R}^{n \times n}$  such that

$$A = U \Sigma X^{-1}, \quad B = V \Lambda X^{-1}, \tag{1.13}$$

where

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I_{n-s} & 0 & 0 \\ 0 & S_A & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \Lambda = [0 \ A_1] = \begin{bmatrix} 0 & S_B \\ 0 & I_{s-t} \end{bmatrix}$$

with

$$S_A = \text{diag}(\alpha_1, \dots, \alpha_t), \quad S_B = \text{diag}(\beta_1, \dots, \beta_t),$$

$$\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_t > 0, \quad 0 < \beta_1 \leq \beta_2 \leq \dots \leq \beta_t, \quad \alpha_i^2 + \beta_i^2 = 1, \quad i = 1, \dots, t,$$

and  $t = \text{rank}(A) + s - n$ . This decomposition is called the GSVD of a matrix pair (see e.g., [13, p. 309], [27]). When  $B = 0$ , the GSVD can reduce to the SVD of  $A$ :

$$A = U \Sigma X^T, \tag{1.14}$$

where  $X$  is orthogonal and  $\Sigma_1 = \text{diag}(\sigma_1, \dots, \sigma_{\text{rank}(A)})$  with  $\sigma_i$  being the  $i$ -th singular value of  $A$ .

## 2 The partial condition number

Let  $L \in \mathbb{R}^{n \times k}$  with  $k \leq n$ . We consider the following function

$$g : \mathbb{R}^{m \times n} \times \mathbb{R}^{s \times n} \times \mathbb{R}^m \times \mathbb{R}^s \rightarrow \mathbb{R}^k$$

$$(A, B, b, d) \rightarrow g(A, B, b, d) = L^T x(A, B, b, d) = L^T (AP)^\dagger b + L^T B_A^\dagger d.$$

From [8,21], it can be seen that the function  $g$  is continuously Fréchet differentiable in a neighborhood of  $(A, B, b, d)$ . Thus, denoting by  $g'$  the Fréchet derivative of  $g$ , and using the chain rules of composition of derivatives or from [8,21], we have

$$g'(A, B, b, d) : \mathbb{R}^{m \times n} \times \mathbb{R}^{s \times n} \times \mathbb{R}^m \times \mathbb{R}^s \rightarrow \mathbb{R}^k$$

$$(\Delta A, \Delta B, \Delta b, \Delta d) \rightarrow g'(A, B, b, d) \circ (\Delta A, \Delta B, \Delta b, \Delta d)$$

$$= L^T ((AP)^T AP)^\dagger (\Delta A)^T r - L^T (AP)^\dagger (\Delta A)x + L^T (AP)^\dagger (\Delta b)$$

$$- L^T ((AP)^T AP)^\dagger (\Delta B)^T (AB_A^\dagger)^T r - L^T B_A^\dagger (\Delta B)x + L^T B_A^\dagger (\Delta d).$$

Here,  $g'(A, B, b, d) \circ (\Delta A, \Delta B, \Delta b, \Delta d)$  denotes that we apply the function  $g'(A, B, b, d)$  to the perturbation variable  $(\Delta A, \Delta B, \Delta b, \Delta d)$  at the point  $(A, B, b, d)$  and  $r = b - Ax$  is called the residual vector. Thus, according to [11,24], we have the absolute normwise condition number of  $g$  at the point  $(A, B, b, d)$  based on the weighted Frobenius norm (1.5):

$$\kappa_{LSE}(A, B, b, d) = \max_{(\alpha_A \Delta A, \alpha_B \Delta B, \alpha_b \Delta b, \alpha_d \Delta d) \neq 0} \frac{\|g'(A, B, b, d) \circ (\Delta A, \Delta B, \Delta b, \Delta d)\|_2}{\|(\alpha_A \Delta A, \alpha_B \Delta B, \alpha_b \Delta b, \alpha_d \Delta d)\|_F}. \tag{2.1}$$

As mentioned in Sect. 1, the condition number  $\kappa_{LSE}(A, B, b, d)$  is called the partial condition number of the LSE problem (1.1) with respect to  $L$ .

In the following, we provide an expression of  $\kappa_{LSE}(A, B, b, d)$ .

**Theorem 2.1** *The partial condition number of the LSE problem (1.1) with respect to  $L$  is*

$$\kappa_{LSE}(A, B, b, d) = \|M_{g'}\|_2, \tag{2.2}$$

where

$$M_{g'} = [M_1, M_2, M_3, M_4] \tag{2.3}$$

with

$$\begin{aligned} M_1 &= \frac{(r^T \otimes (L^T((AP)^T AP)^\dagger)) \Pi_{mn} - x^T \otimes (L^T(AP)^\dagger)}{\alpha_A}, \\ M_2 &= -\frac{\left((r^T AB_A^\dagger) \otimes (L^T((AP)^T AP)^\dagger)\right) \Pi_{sm} + x^T \otimes (L^T B_A^\dagger)}{\alpha_B}, \\ M_3 &= \frac{L^T(AP)^\dagger}{\alpha_b}, \quad M_4 = \frac{L^T B_A^\dagger}{\alpha_d}. \end{aligned}$$

*Proof* Applying the operator  $\text{vec}$  to  $g'(A, B, b, d) \circ (\Delta A, \Delta B, \Delta b, \Delta d)$  and using (1.9) and (1.10), we have

$$\begin{aligned} &g'(A, B, b, d) \circ (\Delta A, \Delta B, \Delta b, \Delta d) = \text{vec}(g'(A, B, b, d) \circ (\Delta A, \Delta B, \Delta b, \Delta d)) \\ &= \left(r^T \otimes (L^T((AP)^T AP)^\dagger)\right) \Pi_{mn} \text{vec}(\Delta A) - \left(x^T \otimes (L^T(AP)^\dagger)\right) \text{vec}(\Delta A) \\ &\quad - \left((r^T AB_A^\dagger) \otimes (L^T((AP)^T AP)^\dagger)\right) \Pi_{sm} \text{vec}(\Delta B) - \left(x^T \otimes (L^T B_A^\dagger)\right) \text{vec}(\Delta B) \\ &\quad + L^T(AP)^\dagger(\Delta b) + L^T B_A^\dagger(\Delta d) \\ &= M_{g'} \begin{bmatrix} \alpha_A \text{vec}(\Delta A) \\ \alpha_B \text{vec}(\Delta B) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix}. \end{aligned}$$

Thus, considering (1.5) and the fact that for any matrix  $C$ ,  $\|C\|_F = \|\text{vec}(C)\|_2$ ,

$$\kappa_{LSE}(A, B, b, d) = \max_{(\alpha_A \Delta A, \alpha_B \Delta B, \alpha_b \Delta b, \alpha_d \Delta d) \neq 0} \frac{\left\| \begin{bmatrix} \alpha_A \text{vec}(\Delta A) \\ \alpha_B \text{vec}(\Delta B) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix} \right\|_2}{\left\| \begin{bmatrix} \alpha_A \text{vec}(\Delta A) \\ \alpha_B \text{vec}(\Delta B) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix} \right\|_2} = \|M_{g'}\|_2. \tag{2.4}$$

□

*Remark 2.1* Setting  $L = I_n$  and  $\alpha_A = \alpha_B = \alpha_b = \alpha_d = 1$  in (2.2), and using the property on the spectral norm that for the matrices  $C$  and  $D$  of suitable orders,  $\| [C, D] \|_2 \leq \| C \|_2 + \| D \|_2$ , we have

$$\begin{aligned} \kappa_{LSE}(A, B, b, d) &\leq \left\| \left( r^T \otimes ((AP)^T AP)^\dagger \right) \Pi_{mn} - x^T \otimes (AP)^\dagger \right\|_2 \\ &\quad + \left\| \left( (r^T AB_A^\dagger) \otimes ((AP)^T AP)^\dagger \right) \Pi_{sm} + x^T \otimes B_A^\dagger \right\|_2 \\ &\quad + \left\| (AP)^\dagger \right\|_2 + \left\| B_A^\dagger \right\|_2 \\ &= \kappa_{LSEup}(A, B, b, d), \end{aligned}$$

where  $\kappa_{LSEup}(A, B, b, d)$  is essentially the same as the upper bound for the normwise condition number of the LSE problem (1.1) obtained in [8]. Note that  $\kappa_{LSEup}(A, B, b, d) \leq 4\kappa_{LSE}(A, B, b, d)$ , which can be verified from a simple fact that  $\| C \|_2 + \| D \|_2 + \| E \|_2 + \| F \|_2 \leq 4 \| [C, D, E, F] \|_2$ . Hence, the bound is a good approximation of the normwise condition number  $\kappa_{LSE}(A, B, b, d)$ . The numerical results given in Sect. 5 confirm this claim.

Note that the expression of the partial condition number  $\kappa_{LSE}(A, B, b, d)$  given in Theorem 2.1 contains Kronecker product. This introduces some large sparse matrices. The following theorem provides a closed formula of  $\kappa_{LSE}(A, B, b, d)$  without Kronecker product.

**Theorem 2.2** *A closed formula of the partial condition number  $\kappa_{LSE}(A, B, b, d)$  is given by*

$$\kappa_{LSE}(A, B, b, d) = \| C \|_2^{1/2}, \tag{2.5}$$

where

$$\begin{aligned} C &= \left( \frac{\| r \|^2}{\alpha_A^2} + \frac{\| r^T AB_A^\dagger \|^2}{\alpha_B^2} \right) L^T ((AP)^T AP)^\dagger)^2 L \\ &\quad + \left( \frac{\| x \|^2}{\alpha_A^2} + \frac{1}{\alpha_b^2} \right) L^T ((AP)^T AP)^\dagger) L \\ &\quad + \left( \frac{\| x \|^2}{\alpha_B^2} + \frac{1}{\alpha_d^2} \right) L^T B_A^\dagger (B_A^\dagger)^T L + \frac{1}{\alpha_B^2} L^T ((AP)^T AP)^\dagger x r^T AB_A^\dagger (B_A^\dagger)^T L \\ &\quad + \frac{1}{\alpha_B^2} L^T B_A^\dagger (B_A^\dagger)^T A^T r x^T ((AP)^T AP)^\dagger L. \end{aligned} \tag{2.6}$$

*Proof* Noting

$$\| M_{g'} \|_2 = \left\| M_{g'} M_{g'}^T \right\|_2^{1/2} = \left\| M_1 M_1^T + M_2 M_2^T + M_3 M_3^T + M_4 M_4^T \right\|_2^{1/2},$$

and

$$M_3M_3^T = \frac{L^T((AP)^T AP)^\dagger L}{\alpha_b^2}, \quad M_4M_4^T = \frac{L^T B_A^\dagger (B_A^\dagger)^T L}{\alpha_d^2}, \tag{2.7}$$

it suffices to obtain the expressions of  $M_1M_1^T$  and  $M_2M_2^T$ .

Let

$$M_{11} = \left( r^T \otimes (L^T((AP)^T AP)^\dagger) \right) \Pi_{mn}, \quad M_{12} = x^T \otimes (L^T(AP)^\dagger).$$

Then

$$M_1M_1^T = \frac{1}{\alpha_A^2} \left( M_{11}M_{11}^T + M_{12}M_{12}^T - M_{11}M_{12}^T - M_{12}M_{11}^T \right).$$

By (1.8) and (1.12), we have

$$\begin{aligned} M_{11}M_{11}^T &= \left( r^T \otimes (L^T((AP)^T AP)^\dagger) \right) \left( r \otimes (((AP)^T AP)^\dagger L) \right) \\ &= \|r\|_2^2 L^T((AP)^T AP)^\dagger{}^2 L, \\ M_{12}M_{12}^T &= \left( x^T \otimes (L^T(AP)^\dagger) \right) \left( x \otimes (((AP)^\dagger)^T L) \right) = \|x\|_2^2 L^T((AP)^T AP)^\dagger L. \end{aligned}$$

Note that

$$\begin{aligned} (AP)^\dagger r &= (AP)^\dagger(b - Ax) = x - B_A^\dagger d - (AP)^\dagger Ax \quad \text{by (1.3)} \\ &= x - B_A^\dagger Bx - (AP)^\dagger Ax = 0. \end{aligned}$$

The last equality in the above equation follows from the GSVD of the matrix pair  $A, B$  in (1.13) and the expressions on  $(AP)^\dagger$  and  $B_A^\dagger$  in Remark 2.3 below. In fact,

$$\begin{aligned} B_A^\dagger B &= X \Lambda^\dagger \Lambda X^{-1} = X \begin{bmatrix} 0 & 0 \\ 0 & I_s \end{bmatrix} X^{-1}, \\ (AP)^\dagger A &= X \begin{bmatrix} I_{n-s} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Sigma X^{-1} = X \begin{bmatrix} I_{n-s} & 0 \\ 0 & 0 \end{bmatrix} X^{-1}, \end{aligned}$$

which mean that  $B_A^\dagger B + (AP)^\dagger A = I_n$  and hence  $x - B_A^\dagger Bx - (AP)^\dagger Ax = 0$ . Thus, by (1.8), (1.11), and (1.12),

$$\begin{aligned} M_{11}M_{12}^T &= \left( r^T \otimes (L^T((AP)^T AP)^\dagger) \right) \left( (((AP)^\dagger)^T L) \otimes x \right) \quad \text{by (1.8) and (1.11)} \\ &= (r^T((AP)^\dagger)^T L) \otimes (L^T((AP)^T AP)^\dagger x) \quad \text{by (1.12)} \\ &= 0 = (M_{12}M_{11}^T)^T. \end{aligned}$$



As a result,

$$M_1 M_1^T = \frac{1}{\alpha_A^2} \left( \|r\|_2^2 L^T ((AP)^T AP)^\dagger)^2 L + \|x\|_2^2 L^T ((AP)^T AP)^\dagger) L \right). \tag{2.8}$$

Now, let

$$M_{21} = \left( (r^T AB_A^\dagger) \otimes (L^T ((AP)^T AP)^\dagger) \right) \Pi_{sm}, \quad M_{22} = x^T \otimes (L^T B_A^\dagger).$$

Then

$$M_2 M_2^T = \frac{1}{\alpha_B^2} \left( M_{21} M_{21}^T + M_{22} M_{22}^T + M_{21} M_{22}^T + M_{22} M_{21}^T \right). \tag{2.9}$$

By (1.8) and (1.12), we get

$$\begin{aligned} M_{21} M_{21}^T &= \left( (r^T AB_A^\dagger) \otimes (L^T ((AP)^T AP)^\dagger) \right) \left( (r^T AB_A^\dagger)^T \otimes (((AP)^T AP)^\dagger L) \right) \\ &= \left\| r^T AB_A^\dagger \right\|_2^2 L^T ((AP)^T AP)^\dagger)^2 L, \end{aligned} \tag{2.10}$$

$$M_{22} M_{22}^T = \left( x^T \otimes (L^T B_A^\dagger) \right) \left( x \otimes ((B_A^\dagger)^T L) \right) = \|x\|_2^2 L^T B_A^\dagger (B_A^\dagger)^T L, \tag{2.11}$$

and by (1.8), (1.11), (1.12), and (1.7), we get

$$\begin{aligned} M_{21} M_{22}^T &= \left( (r^T AB_A^\dagger) \otimes (L^T ((AP)^T AP)^\dagger) \right) \left( ((B_A^\dagger)^T L) \otimes x \right) \text{ by (1.8) and (1.11)} \\ &= (r^T AB_A^\dagger (B_A^\dagger)^T L) \otimes (L^T ((AP)^T AP)^\dagger x) \text{ by (1.12)} \\ &= L^T ((AP)^T AP)^\dagger x r^T AB_A^\dagger (B_A^\dagger)^T L \text{ by (1.7)} \\ &= \left( M_{22} M_{21}^T \right)^T. \end{aligned} \tag{2.12}$$

Substituting (2.10)–(2.12) into (2.9) gives

$$\begin{aligned} M_2 M_2^T &= \frac{1}{\alpha_B^2} \left( \left\| r^T AB_A^\dagger \right\|_2^2 L^T ((AP)^T AP)^\dagger)^2 L + \|x\|_2^2 L^T B_A^\dagger (B_A^\dagger)^T L \right) \\ &\quad + \frac{1}{\alpha_B^2} L^T ((AP)^T AP)^\dagger x r^T AB_A^\dagger (B_A^\dagger)^T L \\ &\quad + \frac{1}{\alpha_B^2} L^T B_A^\dagger (B_A^\dagger)^T A^T r x^T ((AP)^T AP)^\dagger L. \end{aligned} \tag{2.13}$$

From (2.7), (2.8), and (2.13), we have the desired result (2.5). □

*Remark 2.2* When  $B = 0$  and  $d = 0$ , that is, when the LSE problem reduces to the LLS problem (1.4),  $P = I_n$  and  $\text{rank}(A) = n$ . Thus,

$$((AP)^T (AP))^\dagger = (A^T A)^{-1}, \quad B_A^\dagger = 0,$$

and hence

$$\kappa_{LLS}(A, b) = \left\| \frac{\|r\|_2^2}{\alpha_A^2} L^T (A^T A)^{-2} L + \left( \frac{\|x\|_2^2}{\alpha_A^2} + \frac{1}{\alpha_b^2} \right) L^T (A^T A)^{-1} L \right\|_2^{1/2}, \tag{2.14}$$

which is the closed formula of the partial condition number of the LLS problem.

Furthermore, if  $L$  is a column vector, i.e.,  $k = 1$ , then

$$\begin{aligned} \kappa_{LLS}(A, b) &= \left( \frac{\|r\|_2^2}{\alpha_A^2} L^T (A^T A)^{-2} L + \left( \frac{\|x\|_2^2}{\alpha_A^2} + \frac{1}{\alpha_b^2} \right) L^T (A^T A)^{-1} L \right)^{1/2} \\ &= \left( \frac{\|r\|_2^2}{\alpha_A^2} \left\| L^T (A^T A)^{-1} \right\|_2^2 + \left( \frac{\|x\|_2^2}{\alpha_A^2} + \frac{1}{\alpha_b^2} \right) \left\| L^T A^\dagger \right\|_2^2 \right)^{1/2}, \end{aligned} \tag{2.15}$$

which is just the result given in Corollary 1 in [1].

*Remark 2.3* Using the GSVD of the matrix pair  $A, B$  in (1.13), and (3.3), (3.4), and (3.15) in [28], we have

$$\begin{aligned} (AP)^\dagger &= X(\Sigma(I_n - \Lambda^\dagger \Lambda))^\dagger U^T \\ &= X \left( \left[ \begin{array}{ccc} I_{n-s} & 0 & 0 \\ 0 & S_A & 0 \\ 0 & 0 & 0 \end{array} \right] \left( I_n - \left[ \begin{array}{cc} 0 & 0 \\ S_B^{-1} & 0 \\ 0 & I_{s-t} \end{array} \right] \left[ \begin{array}{ccc} 0 & S_B & 0 \\ 0 & 0 & I_{s-t} \end{array} \right] \right) \right)^\dagger U^T \\ &= X \left( \left[ \begin{array}{ccc} I_{n-s} & 0 & 0 \\ 0 & S_A & 0 \\ 0 & 0 & 0 \end{array} \right] \left[ \begin{array}{ccc} I_{n-s} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right] \right)^\dagger U^T = X \left[ \begin{array}{ccc} I_{n-s} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right] U^T \end{aligned}$$

and

$$\begin{aligned} B_A^\dagger &= (I_n - (AP)^\dagger A) X \Lambda^\dagger V^T \\ &= \left( I_n - X \left[ \begin{array}{ccc} I_{n-s} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right] U^T U \left[ \begin{array}{ccc} I_{n-s} & 0 & 0 \\ 0 & S_A & 0 \\ 0 & 0 & 0 \end{array} \right] X^{-1} \right) X \Lambda^\dagger V^T \\ &= X \left[ \begin{array}{ccc} 0 & 0 & 0 \\ 0 & I_t & 0 \\ 0 & 0 & I_{s-t} \end{array} \right] \left[ \begin{array}{cc} 0 & 0 \\ S_B^{-1} & 0 \\ 0 & I_{s-t} \end{array} \right] V^T = X \left[ \begin{array}{cc} 0 & 0 \\ S_B^{-1} & 0 \\ 0 & I_{s-t} \end{array} \right] V^T = X \Lambda^\dagger V^T. \end{aligned}$$

Then

$$(AP)^\dagger \left( (AP)^\dagger \right)^T = X_1 X_1^T, \quad AB_A^\dagger = U \left[ \begin{array}{cc} 0 & 0 \\ S_A S_B^{-1} & 0 \\ 0 & 0 \end{array} \right] V^T = U_2 \left[ \begin{array}{cc} S_A S_B^{-1} & 0 \\ 0 & 0 \end{array} \right] V^T, \tag{2.16}$$

$$B_A^\dagger (B_A^\dagger)^T = X \begin{bmatrix} 0 & 0 & 0 \\ 0 & S_B^{-2} & 0 \\ 0 & 0 & I_{s-t} \end{bmatrix} X^T = X_2 \begin{bmatrix} S_B^{-2} & 0 \\ 0 & I_{s-t} \end{bmatrix} X_2^T, \tag{2.17}$$

$$AB_A^\dagger (B_A^\dagger)^T = U \begin{bmatrix} 0 & 0 & 0 \\ 0 & S_A S_B^{-2} & 0 \\ 0 & 0 & 0 \end{bmatrix} X^T = U_2 \begin{bmatrix} S_A S_B^{-2} & 0 \\ 0 & 0 \end{bmatrix} X_2^T, \tag{2.18}$$

where  $X = [X_1, X_2]$  with  $X_1 \in \mathbb{R}^{n \times (n-s)}$  and  $X_2 \in \mathbb{R}^{n \times s}$ , and  $U = [U_1, U_2]$  with  $U_1 \in \mathbb{R}^{m \times (n-s)}$  and  $U_2 \in \mathbb{R}^{m \times (m-n+s)}$ . Substituting (2.16)–(2.18) into (2.6) yields

$$C = \left( \frac{\|r\|_2^2}{\alpha_A^2} + \frac{\left\| r^T U_2 \begin{bmatrix} S_A S_B^{-1} & 0 \\ 0 & 0 \end{bmatrix} \right\|_2^2}{\alpha_B^2} \right) L^T (X_1 X_1^T)^2 L + L^T (X_1 S_1 X_1^T + X_2 S_2 X_2^T) L \\ + \frac{1}{\alpha_B^2} L^T X_1 X_1^T x r^T U_2 \begin{bmatrix} S_A S_B^{-2} & 0 \\ 0 & 0 \end{bmatrix} X_2^T L + \frac{1}{\alpha_B^2} L^T X_2 \begin{bmatrix} S_A S_B^{-2} & 0 \\ 0 & 0 \end{bmatrix} U_2^T r x^T X_1 X_1^T L \tag{2.19}$$

with

$$S_1 = \left( \frac{\|x\|_2^2}{\alpha_A^2} + \frac{1}{\alpha_b^2} \right) I_{n-s}, \quad S_2 = \left( \frac{\|x\|_2^2}{\alpha_B^2} + \frac{1}{\alpha_d^2} \right) \Lambda_1^{-2}. \tag{2.20}$$

In particular, when  $B = 0$ , the GSVD (1.13) reduces to the SVD of  $A$  (1.14). In this case,  $P = I_n$  and  $\text{rank}(A) = n$ . Hence, we have

$$(AP)^\dagger = X \Sigma^\dagger U^T = X [\Sigma_1^{-1}, 0] U^T, \quad B_A^\dagger = 0,$$

and

$$(AP)^\dagger \left( (AP)^\dagger \right)^T = X \Sigma_1^{-2} X^T.$$

Thus,

$$C = \frac{\|r\|_2^2}{\alpha_A^2} L^T X \Sigma_1^{-4} X^T L + \left( \frac{\|x\|_2^2}{\alpha_A^2} + \frac{1}{\alpha_b^2} \right) L^T X \Sigma_1^{-2} X^T L.$$

As a result, we get a closed formula of the partial condition number of the LLS problem based on the SVD of  $A$ :

$$\kappa_{LLS}(A, b) = \left\| SX^T L \right\|_2, \tag{2.21}$$

where  $S$  is a diagonal matrix with the  $i$ -th diagonal element being

$$S_{ii} = \frac{1}{\sigma_i} \sqrt{\frac{\sigma_i^{-2} \|r\|_2^2 + \|x\|_2^2}{\alpha_A^2} + \frac{1}{\alpha_b^2}}.$$

The closed formula (2.21) is just the one given in Theorem 1 in [1], where it was derived by a different approach.

*Remark 2.4* From (2.19) and (2.5) with setting  $L = I_n$  and  $\alpha_A = \alpha_B = \alpha_b = \alpha_d = 1$  and using (1.13), we have the following upper bound for the normwise condition number of the LSE problem:

$$\kappa_{LSE}(A, B, b, d) \leq (a_1 + a_2 + a_3)^{1/2},$$

where

$$a_1 = \|r\|_2^2 \left(1 + \frac{\alpha_1^2}{\beta_1^2}\right) \|X\|_2^4, \quad a_2 = (1 + \|x\|_2^2) \frac{1}{\beta_1^2} \|X\|_2^2, \quad a_3 = 2 \frac{\alpha_1}{\beta_1^2} \|r\|_2 \|x\|_2 \|X\|_2^3.$$

Moreover, using (2.8) in [29], we obtain  $\|X\|_2 = \frac{1}{\sigma_n}$ , where  $\sigma_n$  is the smallest singular value of  $\begin{bmatrix} A \\ B \end{bmatrix}$ . Hence, the main factors causing the ill-conditioning of LSE problem are  $\|r\|_2$ ,  $\beta_1$ , and  $\sigma_n$ .

### 3 The structured partial condition number

Suppose that  $\mathbb{S}_1 \subseteq \mathbb{R}^{m \times n}$  and  $\mathbb{S}_2 \subseteq \mathbb{R}^{s \times n}$  are two linear subspaces, which consist of two classes of structured matrices, respectively. From [15, 20, 25], we have that if  $A \in \mathbb{S}_1$  and  $B \in \mathbb{S}_2$ , then

$$\text{vec}(A) = \Phi_{\mathbb{S}_1} s_1, \quad \text{vec}(B) = \Phi_{\mathbb{S}_2} s_2, \tag{3.1}$$

where  $\Phi_{\mathbb{S}_1} \in \mathbb{R}^{mn \times k_1}$  and  $\Phi_{\mathbb{S}_2} \in \mathbb{R}^{sn \times k_2}$  are the fixed structure matrices reflecting the structures of  $\mathbb{S}_1$  and  $\mathbb{S}_2$ , respectively, and  $s_1 \in \mathbb{R}^{k_1}$  and  $s_2 \in \mathbb{R}^{k_2}$  are the vectors of the independent parameters in the structured matrices, respectively. Based on the above explanation, the structured perturbations  $\Delta A \in \mathbb{S}_1$  and  $\Delta B \in \mathbb{S}_2$  can be written as

$$\text{vec}(\Delta A) = \Phi_{\mathbb{S}_1}(\Delta s_1), \quad \text{vec}(\Delta B) = \Phi_{\mathbb{S}_2}(\Delta s_2), \tag{3.2}$$

where  $\Delta s_1 \in \mathbb{R}^{k_1}$  and  $\Delta s_2 \in \mathbb{R}^{k_2}$  can be regarded as the perturbations of  $s_1$  and  $s_2$ , respectively.

Now we present the definition of the structured partial condition number of the LSE problem (1.1):

$$\kappa_{LSE}^S(A, B, b, d) = \max_{\substack{(\alpha_A \Delta A, \alpha_B \Delta B, \alpha_b \Delta b, \alpha_d \Delta d) \neq 0 \\ \Delta A \in \mathbb{S}_1, \Delta B \in \mathbb{S}_2}} \frac{\|g'(A, B, b, d) \circ (\Delta A, \Delta B, \Delta b, \Delta d)\|_2}{\|(\alpha_A \Delta A, \alpha_B \Delta B, \alpha_b \Delta b, \alpha_d \Delta d)\|_F},$$

which is a natural variant of the partial condition number in (2.1). From (2.4), it follows that

$$\kappa_{LSE}^S(A, B, b, d) = \max_{\substack{(\alpha_A \Delta A, \alpha_B \Delta B, \alpha_b \Delta b, \alpha_d \Delta d) \neq 0 \\ \Delta A \in \mathbb{S}_1, \Delta B \in \mathbb{S}_2}} \frac{\left\| M_{g'} \begin{bmatrix} \alpha_A \text{vec}(\Delta A) \\ \alpha_B \text{vec}(\Delta B) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix} \right\|_2}{\left\| \begin{bmatrix} \alpha_A \text{vec}(\Delta A) \\ \alpha_B \text{vec}(\Delta B) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix} \right\|_2}. \tag{3.3}$$

Considering (3.2), we have

$$\begin{bmatrix} \text{vec}(\Delta A) \\ \text{vec}(\Delta B) \\ \Delta b \\ \Delta d \end{bmatrix} = \begin{bmatrix} \Phi_{\mathbb{S}_1} & 0 & 0 & 0 \\ 0 & \Phi_{\mathbb{S}_2} & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_s \end{bmatrix} \begin{bmatrix} \Delta s_1 \\ \Delta s_2 \\ \Delta b \\ \Delta d \end{bmatrix}.$$

Substituting the above equation into (3.3) yields

$$\begin{aligned} &\kappa_{LSE}^S(A, B, b, d) \\ &= \max_{(\alpha_A(\Delta s_1), \alpha_B(\Delta s_2), \alpha_b(\Delta b), \alpha_d(\Delta d)) \neq 0} \frac{\left\| M_{g'} \begin{bmatrix} \Phi_{\mathbb{S}_1} & 0 & 0 & 0 \\ 0 & \Phi_{\mathbb{S}_2} & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_s \end{bmatrix} \begin{bmatrix} \alpha_A(\Delta s_1) \\ \alpha_B(\Delta s_2) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix} \right\|_2}{\left\| \begin{bmatrix} \Phi_{\mathbb{S}_1} & 0 & 0 & 0 \\ 0 & \Phi_{\mathbb{S}_2} & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_s \end{bmatrix} \begin{bmatrix} \alpha_A(\Delta s_1) \\ \alpha_B(\Delta s_2) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix} \right\|_2}. \end{aligned} \tag{3.4}$$

Note that

$$\begin{aligned} & \left\| \begin{bmatrix} \Phi_{\mathbb{S}_1} & 0 & 0 & 0 \\ 0 & \Phi_{\mathbb{S}_2} & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_s \end{bmatrix} \begin{bmatrix} \alpha_A(\Delta s_1) \\ \alpha_B(\Delta s_2) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix} \right\|_2 \\ &= \left\| \begin{bmatrix} \alpha_A(\Delta s_1) \\ \alpha_B(\Delta s_2) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix}^T \begin{bmatrix} \Phi_{\mathbb{S}_1}^T \Phi_{\mathbb{S}_1} & 0 & 0 & 0 \\ 0 & \Phi_{\mathbb{S}_2}^T \Phi_{\mathbb{S}_2} & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_s \end{bmatrix} \begin{bmatrix} \alpha_A(\Delta s_1) \\ \alpha_B(\Delta s_2) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix} \right\|_2^{1/2} \end{aligned}$$

and the structured matrices  $\Phi_{\mathbb{S}_1}$  and  $\Phi_{\mathbb{S}_2}$  are column orthogonal [20]. Then

$$\left\| \begin{bmatrix} \Phi_{\mathbb{S}_1} & 0 & 0 & 0 \\ 0 & \Phi_{\mathbb{S}_2} & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_s \end{bmatrix} \begin{bmatrix} \alpha_A(\Delta s_1) \\ \alpha_B(\Delta s_2) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} D_1 & 0 & 0 & 0 \\ 0 & D_2 & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_s \end{bmatrix} \begin{bmatrix} \alpha_A(\Delta s_1) \\ \alpha_B(\Delta s_2) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix} \right\|_2, \tag{3.5}$$

where  $D_1 = \text{diag}(w_1)$  and  $D_2 = \text{diag}(w_2)$  with

$$w_1 = [\|\Phi_{\mathbb{S}_1}(1, :)\|_2, \dots, \|\Phi_{\mathbb{S}_1}(k_1, :)\|_2], \quad w_2 = [\|\Phi_{\mathbb{S}_2}(1, :)\|_2, \dots, \|\Phi_{\mathbb{S}_2}(k_2, :)\|_2].$$

Here, the Matlab notation is used. Combining (3.4) and (3.5) implies

$$\begin{aligned} & \kappa_{LSE}^S(A, B, b, d) \\ &= \frac{\max_{(\alpha_A(\Delta s_1), \alpha_B(\Delta s_2), \alpha_b(\Delta b), \alpha_d(\Delta d)) \neq 0} \left\| M_{g'} \begin{bmatrix} \Phi_{\mathbb{S}_1} D_1^{-1} & 0 & 0 & 0 \\ 0 & \Phi_{\mathbb{S}_2} D_2^{-1} & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_s \end{bmatrix} \begin{bmatrix} D_1 & 0 & 0 & 0 \\ 0 & D_2 & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_s \end{bmatrix} \begin{bmatrix} \alpha_A(\Delta s_1) \\ \alpha_B(\Delta s_2) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix} \right\|_2}{\left\| \begin{bmatrix} D_1 & 0 & 0 & 0 \\ 0 & D_2 & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_s \end{bmatrix} \begin{bmatrix} \alpha_A(\Delta s_1) \\ \alpha_B(\Delta s_2) \\ \alpha_b(\Delta b) \\ \alpha_d(\Delta d) \end{bmatrix} \right\|_2} \end{aligned}$$

Then we can derive the expression of the structured partial condition number of the LSE problem, which is presented in the following theorem.

**Theorem 3.1** *The structured partial condition number of the LSE problem (1.1) with respect to L and the structures  $\mathbb{S}_1$  and  $\mathbb{S}_2$  is*

$$\kappa_{LSE}^S(A, B, b, d) = \left\| M_{g'} \begin{bmatrix} \Phi_{\mathbb{S}_1} D_1^{-1} & 0 & 0 & 0 \\ 0 & \Phi_{\mathbb{S}_2} D_2^{-1} & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_s \end{bmatrix} \right\|_2, \tag{3.6}$$

where  $M_{g'}$  is given in (2.3).

*Remark 3.1* It is easy to verify that

$$\begin{bmatrix} \Phi_{S_1} D_1^{-1} & 0 & 0 & 0 \\ 0 & \Phi_{S_2} D_2^{-1} & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_s \end{bmatrix}$$

is column orthonormal. Thus,

$$\left\| M_{g'} \begin{bmatrix} \Phi_{S_1} D_1^{-1} & 0 & 0 & 0 \\ 0 & \Phi_{S_2} D_2^{-1} & 0 & 0 \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_s \end{bmatrix} \right\|_2 \leq \|M_{g'}\|_2.$$

That is, the structured partial condition number is always tighter than the unstructured one. This fact can also be seen from the definitions of these two condition numbers. As done in [25,31], it is valuable to discuss the ratio between the structured and unstructured partial condition numbers of the LSE problem in detail. We won't go that far in this paper, and only provide a numerical example in Sect. 5 to show that the structured partial condition number is indeed tighter than the unstructured one.

*Remark 3.2* When  $B = 0$  and  $d = 0$ , we have the structured partial condition number of the LLS problem and its upper bound:

$$\kappa_{LLS}^S(A, b) = \left\| \left[ \frac{(r^T \otimes (L^T (A^T A)^{-1})) \Pi_{mn} - x^T \otimes (L^T A^\dagger)}{\alpha_A} \Phi_{S_1} D_1^{-1}, \frac{L^T A^\dagger}{\alpha_b} \right] \right\|_2 \tag{3.7}$$

$$\leq \left\| \left[ \frac{(r^T \otimes (L^T (A^T A)^{-1})) \Pi_{mn} - x^T \otimes (L^T A^\dagger)}{\alpha_A}, \frac{L^T A^\dagger}{\alpha_b} \right] \right\|_2, \tag{3.8}$$

where the upper bound (3.8) is just the unstructured partial condition number of the LLS problem. Here, it should be pointed out that the structured condition number of the LLS problem derived from (3.7) by setting  $L = I_n$  and  $\alpha_A = \alpha_b = 1$  is a little different from the ones in [31] because two additional conditions are added besides the structure requirement in [31].

*Remark 3.3* Similar to Theorem 2.2, we can consider giving the compact expression for the structured condition number (3.6). Unfortunately, the obtained compact expression is very complicated. For example, the terms  $M_{11} M_{11}^T$  and  $M_{12} M_{12}^T$  appearing in the proof of Theorem 2.2 are replaced with the following two ones:

$$M_{11} \Phi_{S_1} D_1^{-2} \Phi_{S_1}^T M_{11}^T = \left( r^T \otimes (L^T ((AP)^T AP)^\dagger) \right) \Pi_{mn} \Phi_{S_1} D_1^{-2} \Phi_{S_1}^T \Pi_{mn}^T \\ \left( r \otimes (((AP)^T AP)^\dagger L) \right).$$

$$M_{12}\Phi_{\mathbb{S}_1}D_1^{-2}\Phi_{\mathbb{S}_1}^T M_{12}^T = \left(x^T \otimes (L^T(AP)^\dagger)\right) \Phi_{\mathbb{S}_1}D_1^{-2}\Phi_{\mathbb{S}_1}^T \left(x \otimes (((AP)^\dagger)^T L)\right).$$

Considering that  $\Phi_{\mathbb{S}_1}$  can be written as

$$\Phi_{\mathbb{S}_1} = [\text{vec}(Z_1), \dots, \text{vec}(Z_{k_1})],$$

where  $Z_1, \dots, Z_{k_1}$  are the basis matrices of the linear subspace  $\mathbb{S}_1$ , and using (1.9) and (1.10), we have

$$\begin{aligned} & \left(r^T \otimes (L^T((AP)^T AP)^\dagger)\right) \Pi_{mn} \Phi_{\mathbb{S}_1} \\ &= \left(r^T \otimes (L^T((AP)^T AP)^\dagger)\right) [\text{vec}(Z_1^T), \dots, \text{vec}(Z_{k_1}^T)] \\ &= \left[L^T((AP)^T AP)^\dagger Z_1^T r, \dots, L^T((AP)^T AP)^\dagger Z_{k_1}^T r\right]. \end{aligned}$$

Thus, writing  $D_1 = \text{diag}(d_1, \dots, d_{k_1})$ , we obtain

$$M_{11}\Phi_{\mathbb{S}_1}D_1^{-2}\Phi_{\mathbb{S}_1}^T M_{11}^T = \sum_{i=1}^{k_1} \frac{1}{d_i^2} L^T((AP)^T AP)^\dagger Z_i^T r r^T Z_i((AP)^T AP)^\dagger L.$$

Similarly, we have

$$M_{12}\Phi_{\mathbb{S}_1}D_1^{-2}\Phi_{\mathbb{S}_1}^T M_{12}^T = \sum_{i=1}^{k_1} \frac{1}{d_i^2} L^T(AP)^\dagger Z_i x x^T Z_i^T ((AP)^\dagger)^T L.$$

Unlike the case in the proof of Theorem 2.2, we cannot combine the two terms

$$M_{11}\Phi_{\mathbb{S}_1}D_1^{-2}\Phi_{\mathbb{S}_1}^T M_{11}^T \text{ and } M_{12}\Phi_{\mathbb{S}_1}D_1^{-2}\Phi_{\mathbb{S}_1}^T M_{12}^T.$$

Worse, we cannot check that the cross term

$$M_{11}\Phi_{\mathbb{S}_1}D_1^{-2}\Phi_{\mathbb{S}_1}^T M_{12}^T = \sum_{i=1}^{k_1} \frac{1}{d_i^2} L^T((AP)^T AP)^\dagger Z_i^T r x^T Z_i^T ((AP)^\dagger)^T L$$

is zero. These facts make the compact expression for the structured condition number (3.6) be very complicated. So, we omit the compact expression here. In addition, we only consider the linear structures of the matrices  $A$  and  $B$  in this section. Similarly, the linear structures of the vectors  $b$  and  $d$  can also be put into the partial condition number. Furthermore, inspired by [9, 20, 26], exploring the structured mixed and componentwise condition numbers of the LSE problem will be interesting. We will investigate this problem in the future research.



### 4 Statistical condition estimates

We first provide a statistical estimate of the partial condition number by using the probabilistic spectral norm estimator. This estimator was proposed by Hochstenbach [16] and can estimate the spectral norm of a matrix reliably. In more detail, the analysis of the estimator in [16] suggests that the spectral norm of a matrix can be contained in a small interval  $[\alpha_1, \alpha_2]$  with high probability, where  $\alpha_1$  is the guaranteed lower bound of the spectral norm of the matrix derived by the famous Lanczos bidiagonalization method [12] and  $\alpha_2$  is the probabilistic upper bound with probability at least  $1 - \epsilon$  with  $\epsilon \ll 1$  derived by finding the largest zero of a polynomial. Meanwhile, we can require  $\alpha_2/\alpha_1 \leq 1 + \delta$  with  $\delta$  being a user-chosen parameter. Based on the above estimator, we can devise Algorithm 1.

---

**Algorithm 1** Probabilistic spectral norm estimator for the partial condition number (2.5)

---

**Input:**  $\epsilon, d$  and matrix  $C$  in (2.6).

**Output:** Probabilistic spectral norm estimator of the partial condition number (2.5):  $\kappa_{PLSE}(A, B, b, d)$ .

1. Get a random starting vector  $v_0$  from  $U(S_{k-1})$ , the uniform distribution over unit sphere  $S_{k-1}$  in  $R^k$ .
2. Compute  $\|C\|_2 \in [\theta_1, \theta_2]$  by probabilistic spectral norm estimator.
  - (a) Determine  $\delta \in [0, \pi/2]$  by  $k, \epsilon$  and  $\int_0^{\arcsin(\delta)} \cos^{k-2}(t) dt = \frac{\epsilon}{2} \int_0^1 t^{\frac{k-3}{2}} (1-t)^{-\frac{1}{2}} dt \left( \epsilon \int_0^{\frac{\pi}{2}} \cos^{k-2}(t) dt \right)$ .
  - (b) **for**  $j = 1, \dots, d$
  - (c)  $u = Cv_j$
  - (d) **if**  $j > 1$
  - (e)  $u = u - \beta_{j-1}u_{j-1}$
  - (f)  $u = u - [u_1, \dots, u_{j-1}](u^T[u_1, \dots, u_{j-1}])^T$
  - (g) **end**
  - (h)  $\alpha_j = \|u\|_2$
  - (i)  $u_j = u/\alpha_j$
  - (j)  $v = C^T u$
  - (k)  $v = v - \alpha_j v_j$
  - (l)  $v = v - [v_1, \dots, v_{j-1}](v^T[v_1, \dots, v_{j-1}])^T$
  - (m)  $v\beta_j = \|v\|_2$
  - (n)  $v_{j+1} = v/\beta_j$
  - (o) **end**
  - (p) Determine the largest singular value  $\theta_1$  of  $B_k$ , and an upper bidiagonal matrix with  $\alpha_i$  on the diagonal and  $\beta_i$  on upper subdiagonal.
  - (q) Determine the probabilistic upper bound  $\theta_2$  for  $\|C\|_2$  with probability  $\geq 1 - \epsilon$  by a Lanczos polynomial (see [16]).
3. Estimate the partial condition number (2.5) by

$$\kappa_{PLSE}(A, B, b, d) = \sqrt{\frac{\theta_1 + \theta_2}{2}}.$$


---

*Remark 4.1* The step 2 of Algorithm 1 is directly taken from [16], and a detailed explanation can be found there. In the practical implementation of Algorithm 1,  $d$  is the dimension of Krylov space and can be automatically determined by the algorithm.

Moreover, as suggested in [16], explicitly forming matrix  $C$  is not necessary because what we really need is the product of a random vector with the matrix  $C$  or  $C^T$ . Hence, some techniques in solving linear system can be employed to reduce the computational burden, especially for large scale problems. Furthermore, it is worthy to point out that Algorithm 1 is also applicable to estimating the partial structured condition number (3.6) since it is also the spectral norm of a matrix.

Now we introduce an alternative approach based on the SSCE method [3, 18] for estimating the normwise condition number of the solution  $x(A, B, b, d)$ . Denote by  $\kappa_{LSEi}(A, B, b, d)$  the normwise condition number of the function  $z_i^T x(A, B, b, d)$ , where  $z_i$ 's are chosen from  $\mathcal{U}(S_{n-1})$  and are orthogonal. Then, from (2.6), we have

$$\begin{aligned} \kappa_{LSEi}^2(A, B, b, d) &= \left( \frac{\|r\|_2^2}{\alpha_A^2} + \frac{\|r^T A B_A^\dagger\|_2^2}{\alpha_B^2} \right) z_i^T ((AP)^T AP)^\dagger z_i \\ &\quad + \left( \frac{\|x\|_2^2}{\alpha_A^2} + \frac{1}{\alpha_b^2} \right) z_i^T ((AP)^T AP)^\dagger z_i \\ &\quad + \left( \frac{\|x\|_2^2}{\alpha_B^2} + \frac{1}{\alpha_d^2} \right) z_i^T B_A^\dagger (B_A^\dagger)^T z_i \\ &\quad + \frac{2}{\alpha_B^2} z_i^T ((AP)^T AP)^\dagger x r^T A B_A^\dagger (B_A^\dagger)^T z_i. \end{aligned} \tag{4.1}$$

The analysis based on SSCE method in [3] shows that

$$\kappa_{SLSE}(A, B, b, d) = \frac{\omega_q}{\omega_n} \sqrt{\sum_{i=1}^q \kappa_{LSEi}^2(A, B, b, d)} \tag{4.2}$$

is a good estimate of the normwise condition number of the LSE problem (1.1). In the above expression,  $\omega_q$  is the Wallis factor with  $\omega_1 = 1$ ,  $\omega_2 = 2/\pi$ , and

$$\omega_q = \begin{cases} \frac{1 \cdot 3 \cdot 5 \cdots (q-2)}{2 \cdot 4 \cdot 6 \cdots (q-1)}, & \text{for } q \text{ odd,} \\ \frac{2}{\pi} \frac{2 \cdot 4 \cdot 6 \cdots (q-2)}{3 \cdot 5 \cdot 7 \cdots (q-1)}, & \text{for } q \text{ even,} \end{cases} \text{ when } q > 2.$$

It can be approximated by

$$\omega_q \approx \sqrt{\frac{2}{\pi(q - \frac{1}{2})}} \tag{4.3}$$

with high accuracy. In summary, we can propose Algorithm 2.

*Remark 4.2* In Algorithm 2,  $\kappa_{LSEi}^2(A, B, b, d)$  is computed by the Eq. (4.1). In practice, the computation of  $\kappa_{LSEi}^2(A, B, b, d)$  should rely on the intermediate results of

---

**Algorithm 2** SSCE method for the normwise condition number of the LSE solution

---

**Input:** Sample size  $q$  and matrix  $C$  in (2.6) with  $L = I_n$ .

**Output:** SSCE estimator of the normwise condition number of the LSE solution:  $\kappa_{SLSE}(A, B, b, d)$ .

1. Generate  $q$  random vectors  $[z_1, \dots, z_q] \rightarrow Z$  from  $\mathcal{U}(S_{n-1})$ .
2. Orthonormalize these vectors using the QR facotization  $[Z, \sim] = QR(Z)$ .
3. For  $i = 1, \dots, q$ , compute  $\kappa_{LSEi}^2(A, B, b, d)$  by (4.1).
4. Approximate  $\omega_q$  and  $\omega_n$  by (4.3).
5. Estimate the normwise condition number by

$$\kappa_{SLSE}(A, B, b, d) = \frac{\omega_q}{\omega_n} \sqrt{\sum_{i=1}^q \kappa_{LSEi}^2(A, B, b, d)}.$$


---

the process for solving the LSE problem to reduce the computational burden. Just as carried out in [3], where the estimate is computed by using the  $R$  factor of QR decomposition, it is better to compute  $\kappa_{LSEi}^2(A, B, b, d)$  through a formula descended from (2.19) instead of (2.6) if we solve the LSE problem by GSVD.

### 4.1 Complexity analysis

Since the main tasks in both Algorithms 1 and 2 are to compute the Moore-Penrose inverses of  $B, AP$  and  $(AP)^T AP$ , and the computation of Moore-Penrose inverse in MATLAB needs SVD, we claim that the computational complexities of these two algorithms are dominated by the computation of SVD. When Golub-Reinsch SVD is used [13], the computational complexities of forming  $B^\dagger, (AP)^\dagger$  and  $((AP)^T AP)^\dagger$  are  $O(5s^2n + 9sn^2 + 9n^3), O(n^2s + 6m^2n + 9mn^2 + 9n^3)$  and  $O(m^2n + mn^2 + 23n^3)$ , respectively. Based on these results, we now present the complexity analysis of Algorithms 1 and 2. We first consider Algorithm 2 whose computational complexity mainly consists of two parts. The first one is from forming  $C$  with  $L = I_n$  which can be given by considering the results mentioned above, and the other one is from the orthonormalization procedure which requires  $O(4n^2q - 2nq^2 + 2/3q^3)$  if Household transformation is used [13]. Summing up the operations on matrix, we obtain the computational complexity of Algorithm 2:  $O(5s^2n + 13sn^2 + 13mn^2 + 5m^2n + 41n^3 + 4n^2q - 2nq^2 + 2/3q^3)$ . Note that in the practical implication of Algorithm 2,  $p$  is usually chosen to be 2 or 3 which is enough for practical application and is much smaller compared with  $n$ . So the computational complexity of  $QR$  decomposition can be omitted. For Algorithm 1, the first part of computational complexity is also to form  $C$ . However, considering that  $d$  is the dimension of Krylov space and can be automatically determined by the stop criterion of iteration  $\epsilon$ , we find that it is not a easy task to give a detailed computational complexity analysis of the rest part of the algorithm in general. In a special case, for example, when  $k$  is given, the computational complexity of the inner loop is about  $O(dk^2)$ . Under this circumstance, the computational complexity of Algorithm 1 is composed of  $O(5s^2n + 13sn^2 + 13mn^2 + 5m^2n + 41n^3 + dk^2)$  and the process of finding  $\delta$  and the probabilistic lower bound.

### 5 Numerical experiments

In this section, we will present two numerical examples to illustrate the reliability of the statistical condition estimates proposed in Sect. 4 and to compare the structured condition number and the unstructured one, respectively. In these two examples, we will set  $\alpha_A = \alpha_B = \alpha_b = \alpha_d = 1$  and the matrix  $L$  be the identity matrix. In addition, the claim in Remark 2.1 will also be verified in Example 5.1.

*Example 5.1* Similar to [23], we construct the random LSE problem through its augmented system (see Equation (3.7) in [10]), and generate the example as follows. Let  $u_1 \in \mathbb{R}^m$ ,  $u_2 \in \mathbb{R}^s$ , and  $v_1, v_2 \in \mathbb{R}^n$  be unit random vectors, and set

$$A = U_1 \begin{bmatrix} D_1 \\ 0 \end{bmatrix} V_1, \quad B = U_2 [D_2 \ 0] V_2, \quad U_i = I_{m(s)} - 2u_i u_i^T, \quad \text{and} \quad V_i = I_n - 2v_i v_i^T,$$

where  $D_1 = n^{-l_1} \text{diag}(n^{l_1}, (n-1)^{l_1}, \dots, 1)$  and  $D_2 = s^{-l_2} \text{diag}(s^{l_2}, (s-1)^{l_2}, \dots, 1)$ . Let the solution  $x$  be  $x = (1, 2^2, \dots, n^2)^T$  and the residual vector  $r = b - Ax$  be a random vector of specified norm. Thus, just letting the lagrange multiplier  $\lambda = -(BB^T)^{-1}BA^T r$ ,  $b = Ax + r$  and  $d = Bx$  gives the desired LSE problem, and it is easy to check that the condition numbers of  $A$  and  $B$  are  $\kappa(A) = n^{l_1}$  and  $\kappa(B) = s^{l_2}$ , respectively. Recall that for any matrix  $C$ , its condition number  $\kappa(C)$  is defined by  $\kappa(C) = \|C\|_2 \|C^\dagger\|_2$ .

In our numerical experiments, we set  $m = 100$ ,  $n = 80$  and  $s = 50$ , and choose the parameters  $\epsilon = 0.001$ ,  $\delta = 0.01$  in Algorithm 1 and  $q = 2$  in Algorithm 2. By varying the condition numbers of  $A$  and  $B$ , and the residual's norm  $\|r\|_2$ , we test the performance of Algorithms 1 and 2. More precisely, for each pair of  $\kappa(A)$  and  $\|r\|_2$  with a fixed  $\kappa(B)$ , 500 random LSE problems are generated and used for the test. The numerical results on mean and variance of the ratios between the statistical condition estimate and the exact condition number defined as

$$r_{sce} = \kappa_{SLE}(A, B, b, d) / \kappa_{LSE}(A, B, b, d),$$

$$r_{pce} = \kappa_{PLE}(A, B, b, d) / \kappa_{LSE}(A, B, b, d)$$

are reported in Table 1. In addition, we also give a numerical verification of the claim given in Remark 2.1. The numerical results of the ratio

$$\kappa_{LSEup}(A, B, b, d) / \kappa_{LSE}(A, B, b, d)$$

are presented in Table 2. In these experiments, we solve the constructed random LSE problems by GSVD, i.e., using the results from Remark 2.3, and compute the exact condition number via (2.19) and (2.5). Here, it should be pointed out that the predetermined exact solution is generally not the same as the computed one. This fact is illustrated in Fig. 1, where, for 50 random LSE problems, we plot the relative error  $\|\hat{x} - x\|_2 / \|x\|_2$  with  $\|r\|_2 = 10^{-1}$ . Moreover, from Fig. 1, we can see that when the coefficient matrices become ill-conditioned, the computed solution by GSVD method gets inaccurate. So we may resort to some other direct methods (see [13, Chapter 12]) or iterative techniques to get more accurate solution for the problem with ill-conditioned

coefficient matrices. The main reasons why we use the GSVD method in this paper are that, with the intermediate results of solving LSE problem by this method, we can compute the condition number via (2.19) and (2.5) conveniently and this method is very effective for the coefficient matrices with small condition numbers. More specifically, from [22], the computational complexity of GSVD is about  $O(8m^2n + 8s^2n + 6mn^2 + 6sn^2 + 8msn + 62/3n^3)$ , which occupies the major computational burden of solving LSE problem. With the intermediate results, we can compute (2.19) with about  $O(4n^3 + ns^2 - 2n^2s)$  flops, and this saves much computational effort compared with calculating (2.6) directly. On the other hand, from the complexity analysis given in 4.1, we need about  $O(4n^3 + ns^2 - 2n^2s + 4n^2q - 2nq^2 + 2/3q^3)$  flops for Algorithm 2, and  $O(4n^3 + ns^2 - 2n^2s + dk^2)$  flops plus the process of finding  $\delta$  and the probabilistic lower bound for Algorithm 1 to estimate the condition number. It is easy to see that the computational burden of estimating the condition number is much smaller than that of computing the exact condition number via GSVD method.

From Table 1, one can easily find that in general both Algorithms 1 and 2 can give reliable estimates of the normwise condition number. In comparison, Algorithm 1 performs more stable since the variances with this algorithm are smaller in most cases. Meanwhile, it should be pointed out that when  $l_1 = l_2 = 0$ , Algorithm 2 may give an inaccurate estimate, i.e., the ratio may be larger than 10. This phenomenon also exists in estimating the normwise condition number of the LLS problem [3]. Although the expression of  $\kappa_{LSE}(A, B, b, d)$  is more complicated than that of the normwise condition number of the LLS problem and the circumstances on these two problems are different, we believe that the underlying reason should be the same; the reader can refer to [3] for a detailed explanation.

Table 2 indicates that  $\kappa_{LSEup}(A, B, b, d)$  is indeed larger than  $\kappa_{LSE}(A, B, b, d)$ . However, the difference between them is not so significant, especially when the condition numbers of the involved matrices are very large. In the latter case, they are the same if we only take four decimal places in our numerical experiments. These numerical results confirm the analysis in Remark 2.1.

*Example 5.2* Let  $A$  and  $B$  be nonsymmetric gaussian random Toeplitz matrices of order  $n = 100$ . This means that the entries of these two matrices are generated from standard normal distribution and the basis matrices for generating the linear space  $\mathbb{S}_1$  or  $\mathbb{S}_2$  are

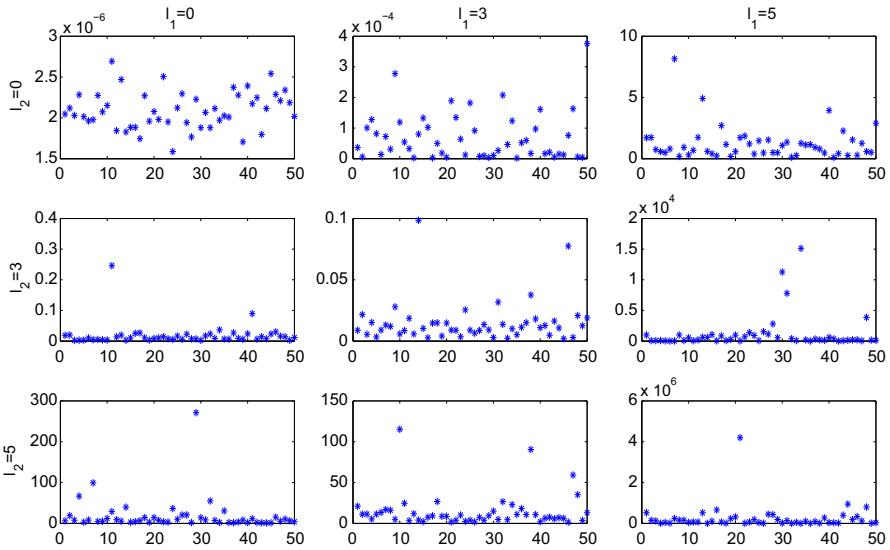
$$\begin{aligned}
 Z_1 &= \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad Z_2 = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \end{bmatrix}, \quad \dots, \quad Z_{n-1} = \begin{bmatrix} 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \\
 Z_n &= \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix},
 \end{aligned}$$

**Table 1** The efficiency of statistical condition estimates with  $\kappa(A) = n^{l_1}$  and  $\kappa(B) = s^{l_2}$

$n^{l_1}$	$s^{l_2}$	$l_2 = 0$		$l_2 = 3$		$l_2 = 5$	
		Mean	Variance	Mean	Variance	Mean	Variance
$\ r\ _2 = 10^{-4}$							
$l_1 = 0$	$r_{ssce}$	1.0296e+001	8.3526e-019	1.4342e+000	2.4351e-001	1.2311e+000	2.6254e-001
	$r_{pce}$	1.0000e+000	7.3038e-012	1.0002e+000	1.5661e-007	1.0000e+000	2.6382e-009
$l_1 = 3$	$r_{ssce}$	1.1726e+000	2.3117e-001	1.3742e+000	2.5506e-001	1.0462e+000	2.6536e-001
	$r_{pce}$	1.0000e+000	1.8003e-012	1.0001e+000	1.2946e-007	1.0000e+000	1.7737e-011
$l_1 = 5$	$r_{ssce}$	1.1114e+000	2.3453e-001	9.9018e-001	2.4687e-001	1.0183e+000	2.7286e-001
	$r_{pce}$	1.0000e+000	6.4792e-012	1.0000e+000	1.9221e-011	1.0000e+000	1.9088e-011
$\ r\ _2 = 10^0$							
$l_1 = 0$	$r_{ssce}$	1.0296e+001	8.1560e-011	1.4357e+000	2.5384e-001	1.2313e+000	2.7339e-001
	$r_{pce}$	1.0000e+000	9.2719e-022	1.0001e+000	1.2321e-007	1.0000e+000	1.4805e-008
$l_1 = 3$	$r_{ssce}$	1.0866e+000	2.3497e-001	1.0552e+000	2.8557e-001	1.0743e+000	2.8316e-001
	$r_{pce}$	1.0000e+000	1.1442e-011	1.0000e+000	1.8651e-011	1.0000e+000	1.7707e-011
$l_1 = 5$	$r_{ssce}$	1.0665e+000	2.4450e-001	1.0337e+000	2.7529e-001	1.0331e+000	2.6395e-001
	$r_{pce}$	1.0000e+000	1.6689e-011	1.0000e+000	1.9375e-011	1.0000e+000	1.9223e-011
$\ r\ _2 = 10^4$							
$l_1 = 0$	$r_{ssce}$	9.4856e+000	6.6559e-003	1.4644e+000	2.7345e-001	1.1969e+000	2.2951e-001
	$r_{pce}$	1.0000e+000	8.0067e-014	1.0001e+000	1.3991e-007	1.0000e+000	2.0099e-008
$l_1 = 3$	$r_{ssce}$	1.0144e+000	2.3900e-001	1.0681e+000	3.1274e-001	1.0387e+000	2.8176e-001
	$r_{pce}$	1.0000e+000	1.7744e-011	1.0000e+000	1.8125e-011	1.0000e+000	1.8783e-011
$l_1 = 5$	$r_{ssce}$	1.0369e+000	2.7324e-001	1.0582e+000	2.7921e-001	1.0344e+000	3.0994e-001
	$r_{pce}$	1.0000e+000	1.7737e-011	1.0000e+000	1.9226e-011	1.0000e+000	1.9394e-011

**Table 2** Ratios of  $\kappa_{LSEup}(A, B, b, d)$  and  $\kappa_{LSE}(A, B, b, d)$  with  $\kappa(A) = n^{l_1}, \kappa(B) = s^{l_2}$  and  $\|r\|_2 = 10^{-1}$

	$l_2 = 0$			$l_2 = 3$			$l_2 = 5$		
	Min	Median	Max	Min	Median	Max	Min	Median	Max
$l_1 = 0$	2.0001	2.0001	2.0001	1.0001	1.0003	1.0008	1.0000	1.0000	1.0000
$l_1 = 3$	1.0199	1.0405	1.2561	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$l_1 = 5$	1.0010	1.2272	1.3192	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000



**Fig. 1** The relative error  $\|\hat{x} - x\|_2 / \|x\|_2$  with  $\|r\|_2 = 10^{-1}$

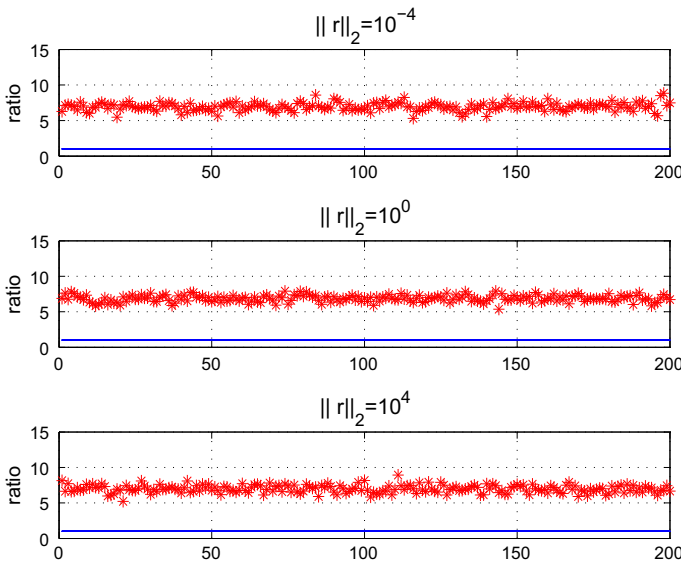
and hence  $\Phi_{S_1} = \Phi_{S_2} = [\text{vec}(Z_1), \dots, \text{vec}(Z_n)]$ . Analogous to Example 5.1, we also let the solution  $x$  be  $x = (1, 2^2, \dots, n^2)^T$  and the residual vector  $r$  be a random vector of specified norm, and obtain the computed solution by GSVD method. However, unlike Example 5.1, it seems impossible to restrict a specific condition number to gaussian random Toeplitz matrices.

In our numerical experiment, for each  $r$ , we test 200 pairs of random Toeplitz matrices  $A$  and  $B$ . The numerical results on the ratio between  $\kappa_{LSE}(A, B, b, d)$  and  $\kappa_{LSE}^S(A, B, b, d)$  defined by

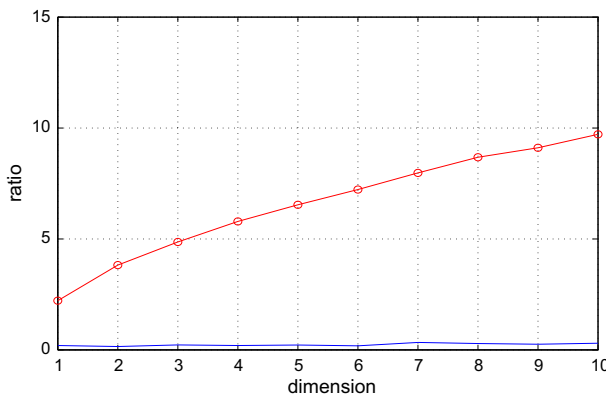
$$ratio = \frac{\kappa_{LSE}(A, B, b, d)}{\kappa_{LSE}^S(A, B, b, d)}$$

are presented in Fig. 2, which confirms the theoretical analysis in Remark 3.1.

From Fig. 2, we also find that there are some points near 10, which means that the unstructured condition number can be 10 times larger than the structured one. Thus, it



**Fig. 2** Comparison of  $\kappa_{LSE}(A, B, b, d)$  and  $\kappa_{LSE}^S(A, B, b, d)$



**Fig. 3** The influence of dimension

may lead to an overestimate when using the unstructured condition number to give error bounds in a structured LSE problem. Moreover, we also note that, for different  $\|r\|_2$ s, the *ratios* seem to follow the same trend gathering in the interval [5, 10]. Whereas, from numerical experiments, we verify that the *ratio* tends to be larger as  $n$  increases. In the numerical experiments, we set  $\|r\|_2 = 1$  and  $n = 20 * i - 10$ ,  $i = 1 : 11$ , and, for every  $n$ , we test 50 LSE problems with random Toeplitz coefficient matrices  $A$  and  $B$ . The numerical results are presented in Fig. 3, where the circle line denotes the mean value of *ratios* and the solid line denotes the corresponding variances. The fact shown in the figure means that the structured condition number has more advantage compared with the unstructured one as the dimensions of coefficient matrices increase.



**Acknowledgements** The authors would like to thank the editor, Froilan M. Dopico, and the anonymous referee for their helpful comments for improving the manuscript. They also would like to acknowledge Prof. Michiel E. Hochstenbach for providing the Matlab program of probabilistic spectral norm estimator. Part of this work was finished when the first author was a visiting scholar at the Department of Mathematics and Statistics of Auburn University from August 2014 to August 2015.

## References

1. Arioli, M., Baboulin, M., Gratton, S.: A partial condition number for linear least squares problems. *SIAM J. Matrix Anal. Appl.* **29**, 413–433 (2007)
2. Baboulin, M., Gratton, S.: A contribution to the conditioning of the total least squares problem. *SIAM J. Matrix Anal. Appl.* **32**, 685–699 (2011)
3. Baboulin, M., Gratton, S., Lacroix, R., Laub, A.J.: Statistical estimates for the conditioning of linear least squares problems. *Lect. Notes Comput. Sci.* **8384**, 124–133 (2014)
4. Barlow, J.L., Nichols, N.K., Plemmons, R.J.: Iterative methods for equality constrained least squares problems. *SIAM J. Sci. Stat. Comput.* **9**, 892–906 (1988)
5. Bergou, E.H., Gratton, S., Tshimanga, J.: The exact condition number of the truncated singular value solution of a linear ill-posed problem. *SIAM J. Matrix Anal. Appl.* **35**, 1073–1085 (2014)
6. Björck, Å.: *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia (1996)
7. Cao, Y., Petzold, L.: A subspace error estimate for linear systems. *SIAM J. Matrix Anal. Appl.* **24**, 787–801 (2003)
8. Cox, A.J., Higham, N.J.: Accuracy and stability of the null space method for solving the equality constrained least squares problem. *BIT* **39**, 34–50 (1999)
9. Cucker, F., Diao, H.: Mixed and componentwise condition numbers for rectangular structured matrices. *Calcolo* **44**, 89–115 (2007)
10. Eldén, L.: Perturbation theory for the least squares problem with equality constraints. *SIAM J. Numer. Anal.* **17**, 338–350 (1980)
11. Geurts, A.J.: A contribution to the theory of condition. *Numer. Math.* **39**, 85–96 (1982)
12. Golub, G.H., Kahan, W.: Calculating the singular values and pseudo-inverse of a matrix. *J. Soc. Ind. Appl. Math. Ser. B Numer. Anal.* **2**, 205–224 (1965)
13. Golub, G., Van Loan, C.F.: *Matrix Computations*, 4th edn. Johns Hopkins University Press, Baltimore (2013)
14. Gratton, S.: On the condition number of linear least squares problems in a weighted Frobenius norm. *BIT* **36**, 523–530 (1996)
15. Higham, D.J., Higham, N.J.: Backward error and condition of structured linear systems. *SIAM J. Matrix Anal. Appl.* **13**, 162–175 (1992)
16. Hochstenbach, M.: Probabilistic upper bounds for the matrix two-norm. *J. Sci. Comput.* **57**, 464–476 (2013)
17. Horn, R.A., Johnson, C.R.: *Topics in Matrix Analysis*. Cambridge UP, New York (1991)
18. Kenney, C., Laub, A.: Small-sample statistical condition estimates for general matrix functions. *SIAM J. Sci. Comput.* **15**, 36–61 (1994)
19. Lawson, C.L., Hanson, R.J.: *Solving Least Squares Problems*. SIAM, Philadelphia (1995)
20. Li, B.Y., Jia, Z.X.: Some results on condition numbers of the scaled total least squares problem. *Linear Algebra Appl.* **435**, 674–686 (2011)
21. Li, H.Y., Wang, S.X., Yang, H.: On mixed and componentwise condition numbers for indefinite least squares problem. *Linear Algebra Appl.* **448**, 104–129 (2014)
22. Paige, C.C., Saunders, M.A.: Towards a generalized singular value decomposition. *SIAM J. Numer. Anal.* **18**(3), 398–405 (1981)
23. Paige, C.C., Saunders, M.A.: LSQR: an algorithm for sparse linear equations and sparse least squares. *ACM Trans. Math. Software* **8**(1), 43–71 (1982)
24. Rice, J.R.: A theory of condition. *SIAM J. Numer. Anal.* **3**, 287–310 (1966)
25. Rump, S.M.: Structured perturbations. Part I: normwise distances. *SIAM J. Matrix Anal. Appl.* **25**, 1–30 (2003)
26. Rump, S.M.: Structured perturbation. Part II: componentwise distances. *SIAM J. Matrix Anal. Appl.* **25**, 31–56 (2003)

27. Van Loan, C.F.: Generalizing the singular value decomposition. *SIAM J. Numer. Anal.* **13**, 76–83 (1976)
28. Wei, M.: Algebraic properties of the rank-deficient equality-constrained and weighted least squares problem. *Linear Algebra Appl.* **161**, 27–43 (1992)
29. Wei, M.: Perturbation theory for rank-deficient equality constrained least squares problem. *SIAM J. Numer. Anal.* **29**, 1462–1481 (1992)
30. Wei, Y., Diao, H., Qiao, S.: Condition number for weighted linear least squares problem and its condition number, Technical report CAS 04–02-SQ, Department of Computing and Software, McMaster University, Hamilton, ON, Canada, (2004)
31. Xu, W., Wei, Y., Qiao, S.: Condition numbers for structured least squares problems. *BIT* **46**, 203–225 (2006)