

# On condition numbers for Moore–Penrose inverse and linear least squares problem involving Kronecker products

Huaian Diao<sup>1</sup>, Weiguo Wang<sup>2</sup>, Yimin Wei<sup>3</sup> and Sanzheng Qiao<sup>4,\*</sup>†

<sup>1</sup>*School of Mathematics and Statistics, Key Laboratory for Applied Statistics of MOE, Northeast Normal University, Chang Chun 130024, China*

<sup>2</sup>*School of Mathematical Sciences, Ocean University of China, Qingdao, 266100, China*

<sup>3</sup>*School of Mathematical Sciences and Shanghai Key Laboratory of Contemporary Applied Mathematics, Fudan University, Shanghai, 200433, China*

<sup>4</sup>*Department of Computing and Software, McMaster University, Hamilton, Ontario L8S 4K1, Canada*

## SUMMARY

In this paper, we investigate the normwise, mixed, and componentwise condition numbers and their upper bounds for the Moore–Penrose inverse of the Kronecker product and more general matrix function compositions involving Kronecker products. We also present the condition numbers and their upper bounds for the associated Kronecker product linear least squares solution with full column rank. In practice, the derived upper bounds for the mixed and componentwise condition numbers for Kronecker product linear least squares solution can be efficiently estimated using the Hager–Higham Algorithm. Copyright © 2012 John Wiley & Sons, Ltd.

Received 26 April 2010; Revised 23 December 2011; Accepted 24 December 2011

KEY WORDS: condition number; Kronecker product; linear least squares; Moore–Penrose inverse; normwise; mixed; componentwise; perturbation

## 1. INTRODUCTION

Condition numbers play an important role in the study of stability of numerical algorithms and complexity analysis. They measure the sensitivity of the solution of a problem to perturbations in the data. Recently, Cucker, Diao, and Wei [1] presented sharp bounds for the condition numbers for the linear least squares (LS) problems. Xu, Wei, and Qiao [2] derived the condition numbers for structured least squares problem. In this paper, we consider the LS problem involving Kronecker products [3–10]

$$\min_v \|(A \otimes B)v - c\|_2, \quad (1.1)$$

where  $A \otimes B$  has full rank. The LS problem involving Kronecker products arises in many applications, such as structured linear total least norm on blind deconvolution problem [11] and constrained LS problem with Kronecker product structure [3]. A well known and important application of the problem is the bivariate problem of surface fitting, where the data points lie on the vertices of a rectangular grid. The least squares approximation can be given by (see e.g., [12, 13]),

$$\min \sum_{(p,q) \in M} [g(x_p, y_q) - f_{p,q}]^2,$$

\*Correspondence to: S. Qiao, Department of Computing and Software, McMaster University, Hamilton, Ontario L8S 4K1, Canada.

†E-mail: qiao@mcmaster.ca

where  $M = \{1, 2, \dots, m_1\} \times \{1, 2, \dots, m_2\}$ ,  $\{(x_p, y_q) \in \mathbb{R}^2 \mid (p, q) \in M\}$  is a finite set of data points with a scalar value  $f_{p,q}$  associated with each  $(x_p, y_q)$ , and  $g$  is a suitable approximation function in a finite dimensional vector space  $V$ . When  $V$  is a tensor product of function spaces, then the least squares problem involving Kronecker products is obtained.

Normwise perturbation analysis, which measures both the input and output data errors by norms, is classical in numerical analysis. Since 1980s, componentwise analysis, which often gives sharper error bounds, has been used. In fact, most error bounds in LAPACK [14] are based on componentwise perturbation analysis. There are two kinds of condition numbers in componentwise analysis: the mixed condition numbers and componentwise condition numbers [15]. The mixed condition numbers use the componentwise error analysis for the input data, whereas the normwise error analysis for the output data. On the other hand, the componentwise condition numbers use the componentwise error analysis for both input and output data. In practice, because of rounding errors and data storage limitation, it is reasonable to measure the input errors componentwise instead of normwise. Moreover, a condition number gives the *worst case* sensitivity measurement. Normwise condition number may overestimate the errors. For example, let

$$A = \begin{bmatrix} 10^{-6} & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad c = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^9$$

in the least squares problem (1.1), then Corollary 4.1 gives the mixed, componentwise condition number and their upper bounds

$$m^{ls}(A \otimes B, c) = m^{ls}(A \otimes B, c)^{\text{upper}} = c^{ls}(A \otimes B, c) = c^{ls}(A \otimes B, c)^{\text{upper}} = 3,$$

and the normwise condition number

$$\kappa^{ls}(A \otimes B, c) = 4.8990e + 006.$$

The mixed and componentwise condition numbers  $m^{ls}(A \otimes B, c)$  and  $c^{ls}(A \otimes B, c)$  are much smaller than the normwise condition number  $\kappa^{ls}(A \otimes B, c)$ . Consequently, the perturbation bounds based on the mixed and componentwise condition numbers are more effective and sharper than those based on the normwise condition number.

We first give a brief review of some previous work on the normwise, mixed, and componentwise perturbation analysis for linear system and LS. Skeel [16] introduced the mixed perturbation analysis of nonsingular linear systems of equations, a mixed error analysis of Gaussian elimination, and obtained the mixed condition number for the solution of nonsingular linear systems. In [17], Rohn derived a new relative condition number measuring the perturbations in both the input and output componentwisely, which was named as componentwise condition number. In [18] and [19], tight upper bounds for the mixed condition numbers for LS problem are given. For structured componentwise perturbation, Rump [20] concentrated on the linear systems with some special structures, such as Toeplitz and Hankel. Cucker and Diau [21] derived an exact expression for LS problems with structures. Arioli *et al.* [22], using normwise perturbation of inputs, obtained expressions of condition numbers for some components of the LS solution. In [23], using dual techniques, the exact expression of the linear functional of LS solution is given.

In this paper, we investigate the normwise, componentwise, and mixed condition numbers for a Kronecker product linear least squares (KPLS) problem. By applying the chain rule to derive the corresponding Fréchet derivatives and then take the appropriate norm, we derive the exact expressions of the condition numbers, which are useful for hybrid symbolic-numeric computations [24]. When the problem size is small, we can obtain the exact condition numbers by using symbolic computations. Moreover, we also derive the corresponding upper bounds. Thus, when the problem size is large, we can apply fast algorithms for solving least squares problems [25] and our upper bounds to numerically estimate the condition numbers. We show that using the Hager–Higham condition estimation algorithm [26, 27], we can efficiently estimate the upper bounds for mixed and componentwise condition numbers, which can be used to estimate the accuracy of the computed solution and the conditioning of the problem.

This paper is organized as follows. In Section 2, we review some concepts related to this paper. In Section 3, we investigate the normwise, mixed, and componentwise condition numbers and their upper bounds for the Moore–Penrose inverse of a Kronecker product. In Section 4, the condition numbers for the associated LS problem involving Kronecker products are studied. In Section 5, a numerical example is provided. Finally, in Section 6, concluding remarks for the future research are presented.

Throughout this paper,  $\mathbb{R}_r^{m \times n}$  denotes the set of real  $m \times n$  matrices of rank  $r$ . For a matrix  $A \in \mathbb{R}^{m \times n}$ ,  $A^T$  is the transpose of  $A$ ;  $\text{rank}(A)$  denotes the rank of  $A$ ;  $\|A\|_2$  is the spectral norm of  $A$ , and  $\|A\|_F$  is the Frobenius norm of  $A$ . The identity matrix of order  $n$  is denoted by  $I_n$ . For a vector  $a$ ,  $\|a\|_\infty$  is its  $\infty$ -norm, and  $\|a\|_1$  is the 1-norm. For any matrix  $A = [a_1 \ a_2 \ \cdots \ a_n] \in \mathbb{R}^{m \times n}$ , we define  $\text{vec}(A) \in \mathbb{R}^{mn}$  by  $\text{vec}(A) = [a_1^T \ a_2^T \ \cdots \ a_n^T]^T$ , by stacking the columns of  $A$ . For a vector  $a$ ,  $|a|$  is the vector whose components are the absolute values of the corresponding components of  $a$ . Also for two vectors  $a, b \in \mathbb{R}^n$ ,  $a \leq b$  means  $a_i \leq b_i$  with  $a_i$  being the  $i$ th component of  $a$ . If  $a \in \mathbb{R}^n$ , then we denote by  $D_a = \text{diag}(a_1, a_2, \dots, a_n)$  the  $n \times n$  diagonal matrix with  $a_1, a_2, \dots, a_n$  in its diagonal. In general,  $D_A = D_{\text{vec}(A)}$ , where  $A$  is a matrix.

## 2. PRELIMINARIES

In this section, we review some concepts and properties that are useful for this paper. First, we define a componentwise vector division. For any vectors  $a, b \in \mathbb{R}^n$ , we define the componentwise division  $\frac{a}{b} = [c_1 \ c_2 \ \cdots \ c_n]^T$  by

$$c_i = \begin{cases} a_i/b_i & \text{if } b_i \neq 0, \\ 0 & \text{if } a_i = b_i = 0, \\ \infty & \text{otherwise.} \end{cases}$$

Using the aforementioned definition, we then define a distance function. Let  $c = \frac{a-b}{b}$ , then the distance  $d(a, b) = \|c\|_\infty = \max_i(|c_i|)$ . In the rest of this paper, we will only consider vector pairs  $(a, b)$  satisfying  $d(a, b) < \infty$ . For matrices, we have the matrix componentwise division  $\frac{A}{B} = \frac{\text{vec}(A)}{\text{vec}(B)}$ . Accordingly, a matrix distance function  $d(A, B) = d(\text{vec}(A), \text{vec}(B))$ . Also, we can define a matrix norm  $\|A\|_{\max} := \|\text{vec}(A)\|_\infty$ , which equals  $\max_{i,j}(|a_{ij}|)$ .

A general theory of condition numbers can be found in [28]. Gohberg and Koltracht [15] gave a useful lemma for exact expressions of the condition numbers. Before presenting their lemma, which relates the condition numbers to the Fréchet derivative, we recall the definition of the Fréchet derivative of a function.

### Definition 2.1

Let  $X$  and  $Y$  be Banach spaces and  $D \subset X$  be an open subset of  $X$ . A function  $L : D \rightarrow Y$  is called Fréchet differentiable at  $a \in D$ , if there exists a bounded linear operator  $F_a : D \rightarrow Y$  such that

$$\lim_{t \rightarrow 0} \frac{\|L(a+t) - L(a) - F_a(t)\|_Y}{\|t\|_X} = 0,$$

where  $\|\cdot\|_X$  and  $\|\cdot\|_Y$  are norms defined in  $X$  and  $Y$ , respectively. Then, the linear operator  $F_a$  is called the Fréchet derivative of  $L$  at  $a$ .

### Lemma 2.1 ([15])

- (a) Let  $F : \mathbb{R}^p \rightarrow \mathbb{R}^q$  be Fréchet differentiable at  $a$  and a continuous mapping defined on an open set  $S_F \subset \mathbb{R}^p$  such that  $0 \notin S_F$ . For a given vector  $a \in S_F$  such that  $F(a) \neq 0$  and the ball  $B(a, \epsilon) = \{x \in \mathbb{R}^p \mid \|x - a\|_2 \leq \epsilon\|a\|_2\}$  for a sufficiently small  $\epsilon > 0$ , the normwise condition number of the mapping  $F$  at the point  $a$  is

$$\kappa(F, a) = \lim_{\epsilon \rightarrow 0} \sup_{\substack{x \in B(a, \epsilon) \\ x \neq a}} \frac{\|F(x) - F(a)\|_2 / \|F(a)\|_2}{\|x - a\|_2 / \|a\|_2} = \frac{\|F'(a)\|_2 \|a\|_2}{\|F(a)\|_2},$$

where  $F'(a)$  denotes the Fréchet derivative of  $F$  at  $a$ .

- (b) Under the assumptions in (a), let  $B^0(a, \epsilon) = \{x \mid |x_i - a_i| \leq \epsilon |a_i|, i = 1, 2, \dots, p\} \subset S_F$ ; the mixed condition number of  $F$  at  $a$  becomes

$$m(F, a) = \lim_{\epsilon \rightarrow 0} \sup_{\substack{x \in B^0(a, \epsilon) \\ x \neq a}} \frac{\|F(x) - F(a)\|_\infty}{\|F(a)\|_\infty} \frac{1}{d(x, a)} = \frac{\|F'(a)D_a\|_\infty}{\|F(a)\|_\infty},$$

where  $D_a = \text{diag}(a_1, a_2, \dots, a_p)$ .

- (c) Suppose  $F(a) = [f_1(a) \ f_2(a) \ \dots \ f_q(a)]^T$  such that  $f_j(a) \neq 0$ , for  $j = 1, 2, \dots, q$ . Then, the componentwise condition number of  $F$  at  $a$  is

$$c(F, a) = \lim_{\epsilon \rightarrow 0} \sup_{\substack{x \in B^0(a, \epsilon) \\ x \neq a}} \frac{d(F(x), F(a))}{d(x, a)} = \|D_{F(a)}^{-1} F'(a) D_a\|_\infty,$$

where  $D_{F(a)} = \text{diag}(f_1(a), f_2(a), \dots, f_p(a))$ .

*Remark 2.1*

(a) It is easy to see that  $B^0(a, \epsilon) \subseteq B(a, \epsilon)$ . Because the condition numbers are the limits of the supremum over some sets, the mixed condition numbers must be smaller than the normwise condition number. If there is small component in the output data, then the componentwise condition number may be bigger than the normwise condition number.

(b) From the definitions of condition numbers, we see that they can give a posteriori error estimations for the computed solution  $F(x)$ . For example,

$$\begin{aligned} \frac{\|F(x) - F(a)\|_2}{\|F(a)\|_2} &\lesssim \kappa(F, a) \frac{\|x - a\|_2}{\|a\|_2} \leq \kappa(F, a)\epsilon, \\ \frac{\|F(x) - F(a)\|_\infty}{\|F(a)\|_\infty} &\lesssim m(F, a)d(x, a) \leq m(F, a)\epsilon, \\ \left\| \frac{F(x) - F(a)}{F(a)} \right\|_\infty &= d(F(x), F(a)) \lesssim c(F, a)d(x, a) \leq c(F, a)\epsilon, \end{aligned}$$

for all  $x \in B^0(a, \epsilon)$ , can be used to estimate the accuracy of the computed solution  $F(x)$ .

For  $A \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{p \times q}$ , their Kronecker product  $A \otimes B \in \mathbb{R}^{mp \times nq}$  is defined by  $A \otimes B = [a_{ij} B]$  [29]. The following results can be found in [29] and [30]:

$$\text{vec}(AXB) = (B^T \otimes A)\text{vec}(X), \tag{2.1}$$

$$\text{vec}(A \otimes B) = (I_n \otimes K_{qm} \otimes I_p) (\text{vec}(A) \otimes \text{vec}(B)), \tag{2.2}$$

$$a \otimes b = \text{vec}(ba^T), \tag{2.3}$$

where  $K_{mn}$  is the commutation matrix defined by

$$K_{mn} = \sum_{i=1}^m \sum_{j=1}^n E_{ij}(m \times n) \otimes E_{ji}(n \times m), \tag{2.4}$$

where  $E_{ij}(m \times n) = e_i^{(m)}(e_j^{(n)})^T \in \mathbb{R}^{m \times n}$  denotes the  $(i, j)$ th elementary matrix and  $e_i^{(m)}$  is the  $i$ th column vector of the  $m$ -by- $m$  identity matrix  $I_m$ . For  $C \in \mathbb{R}^{m \times n}$  and  $y \in \mathbb{R}^n$ , the following relations hold [29, 30]:  $K_{nm}\text{vec}(C) = \text{vec}(C^T)$ ;  $K_{mn}(y \otimes C) = C \otimes y$ , and  $K_{mn}^T = K_{nm}$ .

For an  $m$ -by- $n$  matrix  $A$ , its Moore–Penrose inverse [31], denoted by  $A^\dagger \in \mathbb{R}^{n \times m}$ , is the unique matrix satisfying  $AA^\dagger A = A$ ,  $A^\dagger AA^\dagger = A^\dagger$ ,  $(AA^\dagger)^T = AA^\dagger$ , and  $(A^\dagger A)^T = A^\dagger A$ . The solution of the LS problem  $\min_u \|Au - b\|_2$  can be obtained by  $A^\dagger b$ .

### 3. CONDITION NUMBERS FOR THE MOORE–PENROSE INVERSE OF A KRONECKER PRODUCT

In this section, we will consider the normwise, mixed, and componentwise condition numbers for the Moore–Penrose inverse of a Kronecker product  $A \otimes B$ , which plays an important role in the associated least squares problem  $\min_{v \in \mathbb{R}^{nq}} \|(A \otimes B)v - c\|_2$ .

We begin with the general case of the  $s \times t$  matrix function [32],

$$f(X) = \begin{bmatrix} f_{11}(X) & f_{12}(X) & \cdots & f_{1t}(X) \\ f_{21}(X) & f_{22}(X) & \cdots & f_{2t}(X) \\ \vdots & \vdots & \ddots & \vdots \\ f_{s1}(X) & f_{s2}(X) & \cdots & f_{st}(X) \end{bmatrix}$$

of a matrix  $X$ . In particular, let  $A \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{p \times q}$  and set  $X = A \otimes B$ , we obtain the functional composition involving Kronecker products,

$$f(A \otimes B) = \begin{bmatrix} f_{11}(A \otimes B) & f_{12}(A \otimes B) & \cdots & f_{1t}(A \otimes B) \\ f_{21}(A \otimes B) & f_{22}(A \otimes B) & \cdots & f_{2t}(A \otimes B) \\ \vdots & \vdots & \ddots & \vdots \\ f_{s1}(A \otimes B) & f_{s2}(A \otimes B) & \cdots & f_{st}(A \otimes B) \end{bmatrix}. \tag{3.1}$$

Next, we extend the three types of the condition numbers in [1] to the functional composition  $f(A \otimes B)$  (3.1). We can define the normwise condition number for its Moore–Penrose inverse,

$$\kappa(f^\dagger(A \otimes B)) := \lim_{\epsilon \rightarrow 0} \sup_{\substack{\sqrt{\|\Delta A\|_F^2 + \|\Delta B\|_F^2} \\ \leq \epsilon \sqrt{\|A\|_F^2 + \|B\|_F^2}}} \frac{\|f^\dagger((A + \Delta A) \otimes (B + \Delta B)) - f^\dagger(A \otimes B)\|_F}{\epsilon \|f^\dagger(A \otimes B)\|_F},$$

the mixed condition number,

$$m(f^\dagger(A \otimes B)) := \lim_{\epsilon \rightarrow 0} \sup_{\substack{\|\frac{\Delta A}{A}\|_\infty \leq \epsilon \\ \|\frac{\Delta B}{B}\|_\infty \leq \epsilon}} \frac{\|\text{vec}(f^\dagger((A + \Delta A) \otimes (B + \Delta B)) - f^\dagger(A \otimes B))\|_\infty}{\epsilon \|\text{vec}(f^\dagger(A \otimes B))\|_\infty},$$

and the componentwise condition number,

$$c(f^\dagger(A \otimes B)) := \lim_{\epsilon \rightarrow 0} \sup_{\substack{\|\frac{\Delta A}{A}\|_\infty \leq \epsilon \\ \|\frac{\Delta B}{B}\|_\infty \leq \epsilon}} \frac{1}{\epsilon} \left\| \frac{f^\dagger((A + \Delta A) \otimes (B + \Delta B)) - f^\dagger(A \otimes B)}{f^\dagger(A \otimes B)} \right\|_\infty.$$

*Remark 3.1*

In the aforementioned definitions of the condition numbers, we assume that both  $A$  and  $B$  are nonzero matrices.

The matrix of partial derivatives is referred as the Jacobian matrix of  $f(X)$  at  $X$  [30, 33]. Suppose now that each of  $f_1, f_2, \dots, f_m$  is a real-valued differentiable function of the same  $n \times 1$  vector  $x = [x_1 \ x_2 \ \cdots \ x_n]^T$ . Then, we denote  $f(x) = [f_1(x) \ f_2(x) \ \cdots \ f_m(x)]^T$ ; the Jacobian matrix of  $f$  at  $x$  is given by

$$\frac{\partial f(x)}{\partial x^T} = \begin{bmatrix} \frac{\partial}{\partial x_1} f_1(x) & \frac{\partial}{\partial x_2} f_1(x) & \cdots & \frac{\partial}{\partial x_n} f_1(x) \\ \frac{\partial}{\partial x_1} f_2(x) & \frac{\partial}{\partial x_2} f_2(x) & \cdots & \frac{\partial}{\partial x_n} f_2(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} f_m(x) & \frac{\partial}{\partial x_2} f_m(x) & \cdots & \frac{\partial}{\partial x_n} f_m(x) \end{bmatrix},$$

which is just the Fréchet derivative of  $f$  at  $x$ . Before deriving the Fréchet derivative of the vectorized form of the mapping  $A \otimes B \mapsto \mathbb{f}^\dagger(A \otimes B)$ , we quote the following lemma regarding the derivatives of the Moore–Penrose inverse.

*Lemma 3.1 (Theorem 8.3 [30])*

If  $X$  is an  $m \times n$  matrix of full rank and  $X^\dagger$  denotes its Moore–Penrose inverse, then the derivative  $dX^\dagger$  takes the form

$$dX^\dagger = (I_n - X^\dagger X)(dX^T)X^{\dagger T}X^\dagger + X^\dagger X^{\dagger T}(dX^T)(I_m - XX^\dagger) - X^\dagger(dX)X^\dagger$$

and the Jacobian matrix takes the form

$$\frac{\partial \text{vec}(X^\dagger)}{\partial \text{vec}(X)^T} = \left\{ X^{\dagger T}X^\dagger \otimes (I_n - X^\dagger X) + (I_m - XX^\dagger) \otimes X^\dagger X^{\dagger T} \right\} K_{mn} - (X^{\dagger T} \otimes X^\dagger),$$

where  $K_{mn}$  is defined in (2.4).

Now, we present the Fréchet derivative of the mapping  $(\text{vec}(A), \text{vec}(B)) \mapsto \text{vec}(\mathbb{f}^\dagger(A \otimes B))$ . For simplicity,  $\mathbb{f}^\dagger$  denotes  $\mathbb{f}^\dagger(A \otimes B)$ .

*Lemma 3.2*

Let  $A \in \mathbb{R}^{m \times n}$ ,  $B \in \mathbb{R}^{p \times q}$ , and the mapping  $\psi : \mathbb{R}^{mn} \times \mathbb{R}^{pq} \mapsto \mathbb{R}^{st}$  be  $\psi(a, b) = \text{vec}(\mathbb{f}^\dagger(A \otimes B))$ , where  $a = \text{vec}(A)$ ,  $b = \text{vec}(B)$ . If  $\mathbb{f}$  is continuously differentiable at  $A \otimes B$  and  $\mathbb{f}(A \otimes B)$  has full rank, then  $\psi$  is continuous and Fréchet differentiable, and its derivative is of the form

$$\psi'(a, b) = [\phi_{(A,B)} L_B \quad \phi_{(A,B)} L_A],$$

where

$$\phi_{(A,B)} = \left\{ [\mathbb{f}^{\dagger T} \mathbb{f}^\dagger \otimes (I_t - \mathbb{f}^\dagger \mathbb{f}) + (I_s - \mathbb{f} \mathbb{f}^\dagger) \otimes \mathbb{f}^\dagger \mathbb{f}^{\dagger T}] K_{st} - (\mathbb{f}^{\dagger T} \otimes \mathbb{f}^\dagger) \right\} \left( \frac{\partial \mathbb{f}(A \otimes B)}{\partial \text{vec}(A \otimes B)^T} \right) \tilde{K}$$

and

$$L_A = \text{vec}(A) \otimes I_{pq}, \quad L_B = (I_{mn} \otimes \text{vec}(B)) K_{mn}, \quad \tilde{K} = I_n \otimes K_{qm} \otimes I_p.$$

*Proof*

First, we prove that  $\psi$  is differentiable. Because  $\mathbb{f}(A \otimes B)$  has full rank, then from Lemma 3.1, we know that  $\mathbb{f}^\dagger(A \otimes B)$  is differentiable at  $\mathbb{f}(A \otimes B)$ . Also, because  $\mathbb{f}$  is differentiable at  $A \otimes B$ , we can conclude that  $\psi = \text{vec} \circ \mathbb{f}^\dagger$  is differentiable with respect to  $[(\text{vec}(A))^T (\text{vec}(B))^T]^T$ .

Differentiating  $\mathbb{f}(A \otimes B)$ , it follows from Lemma 3.1 that

$$\begin{aligned} d(\mathbb{f}^\dagger(A \otimes B)) &= (I_t - \mathbb{f}^\dagger \mathbb{f}) \left( d(\mathbb{f}(A \otimes B))^T \right) \mathbb{f}^{\dagger T} \mathbb{f}^\dagger \\ &\quad + \mathbb{f}^\dagger \mathbb{f}^{\dagger T} \left( d(\mathbb{f}(A \otimes B))^T \right) (I_s - \mathbb{f} \mathbb{f}^\dagger) - \mathbb{f}^\dagger (d(\mathbb{f}(A \otimes B))) \mathbb{f}^\dagger \end{aligned}$$

and

$$d(\text{vec}(\mathbb{f}^\dagger(A \otimes B))) = \frac{\left\{ [\mathbb{f}^{\dagger T} \mathbb{f}^\dagger \otimes (I_t - \mathbb{f}^\dagger \mathbb{f}) + (I_s - \mathbb{f} \mathbb{f}^\dagger) \otimes \mathbb{f}^\dagger \mathbb{f}^{\dagger T}] K_{st} - (\mathbb{f}^{\dagger T} \otimes \mathbb{f}^\dagger) \right\} \text{vec}(d(\mathbb{f}(A \otimes B)))}{\text{vec}(d(\mathbb{f}(A \otimes B)))}.$$

Using (2.1), (2.2), and (2.3), we have

$$\begin{aligned} \text{vec}(d(\mathbb{f}(A \otimes B))) &= d(\text{vec}(\mathbb{f}(A \otimes B))) = \left( \frac{\partial \mathbb{f}(A \otimes B)}{\partial \text{vec}(A \otimes B)^T} \right) \{ \text{vec}(dA \otimes B) + \text{vec}(A \otimes dB) \} \\ &= \left( \frac{\partial \mathbb{f}(A \otimes B)}{\partial \text{vec}(A \otimes B)^T} \right) (I_n \otimes K_{qm} \otimes I_p) \{ \text{vec}(dA) \otimes \text{vec}(B) + \text{vec}(A) \otimes \text{vec}(dB) \} \\ &= \left( \frac{\partial \mathbb{f}(A \otimes B)}{\partial \text{vec}(A \otimes B)^T} \right) (I_n \otimes K_{qm} \otimes I_p) \{ \text{vec}(\text{vec}(B)(\text{vec}(dA))^T) + \text{vec}(\text{vec}(dB)(\text{vec}(A))^T) \} \\ &= \left( \frac{\partial \mathbb{f}(A \otimes B)}{\partial \text{vec}(A \otimes B)^T} \right) (I_n \otimes K_{qm} \otimes I_p) \{ (I_{mn} \otimes \text{vec}(B)) K_{mn} \text{vec}(dA) + (\text{vec}(A) \otimes I_{pq}) \text{vec}(dB) \}. \end{aligned}$$

It follows from the definitions of  $\phi_{(A,B)}$ ,  $L_A$ , and  $L_B$  that

$$\begin{aligned} d\left(\text{vec}\left(\mathbb{f}^\dagger(A \otimes B)\right)\right) &= \phi_{(A,B)}(I_{mn} \otimes \text{vec}(B))K_{mn}\text{vec}(dA) + \phi_{(A,B)}(\text{vec}(A) \otimes I_{pq})\text{vec}(dB) \\ &= [\phi_{(A,B)}L_B \quad \phi_{(A,B)}L_A] \begin{bmatrix} d(\text{vec}(A)) \\ d(\text{vec}(B)) \end{bmatrix}. \end{aligned}$$

Then, the Fréchet derivative is found to be  $\psi'(a, b) = [\phi_{(A,B)}L_B \quad \phi_{(A,B)}L_A]$ .  $\square$

From the aforementioned Fréchet derivative, we have the following theorem providing explicit expressions of the condition numbers for the Moore–Penrose inverse of the matrix function of Kronecker products, which will be useful in error free symbolic computation.

*Theorem 3.1*

Let  $A \in \mathbb{R}^{m \times n}$ ,  $B \in \mathbb{R}^{p \times q}$  and the mapping  $\psi(a, b) = \text{vec}(\mathbb{f}^\dagger(A \otimes B))$ , where  $a = \text{vec}(A)$ ,  $b = \text{vec}(B)$ . If  $\mathbb{f}$  is continuously differentiable at  $A \otimes B$  and  $\mathbb{f}(A \otimes B)$  has full rank, we have

$$\begin{aligned} \kappa(\mathbb{f}^\dagger(A \otimes B)) &= \frac{\|[\phi_{(A,B)}L_B \quad \phi_{(A,B)}L_A]\|_2 \sqrt{\|A\|_F^2 + \|B\|_F^2}}{\|\mathbb{f}^\dagger(A \otimes B)\|_F}, \\ m(\mathbb{f}^\dagger(A \otimes B)) &= \frac{\|\phi_{(A,B)}L_B|\text{vec}(|A|) + \phi_{(A,B)}L_A|\text{vec}(|B|)\|_\infty}{\|\text{vec}(\mathbb{f}^\dagger(A \otimes B))\|_\infty}. \end{aligned}$$

Furthermore, if there is no zero element in  $\mathbb{f}^\dagger(A \otimes B)$ , then we obtain

$$c(\mathbb{f}^\dagger(A \otimes B)) = \left\| \frac{|\phi_{(A,B)}L_B|\text{vec}(|A|) + |\phi_{(A,B)}L_A|\text{vec}(|B|)}{\text{vec}(\mathbb{f}^\dagger(A \otimes B))} \right\|_\infty.$$

*Proof*

By Lemmas 2.1 and 3.2, noting that  $a = \text{vec}(A)$ ,  $b = \text{vec}(B)$ , we obtain the normwise condition number for  $\mathbb{f}^\dagger(A \otimes B)$ :

$$\begin{aligned} \kappa(\mathbb{f}^\dagger(A \otimes B)) &= \frac{\|\psi'(a, b)\|_2 \|(a^T, b^T)\|_2}{\|\psi(a, b)\|_2} = \frac{\|[\phi_{(A,B)}L_B \quad \phi_{(A,B)}L_A]\|_2 \sqrt{\|a\|_2^2 + \|b\|_2^2}}{\|\text{vec}(\mathbb{f}^\dagger(A \otimes B))\|_2} \\ &= \frac{\|[\phi_{(A,B)}L_B \quad \phi_{(A,B)}L_A]\|_2 \sqrt{\|A\|_F^2 + \|B\|_F^2}}{\|\mathbb{f}^\dagger(A \otimes B)\|_F}. \end{aligned}$$

Recalling that  $D_A = D_{\text{vec}(A)}$  and denoting  $D_{a,b} = \begin{bmatrix} D_A & \\ & D_B \end{bmatrix}$ , we obtain the following form of the mixed condition number for the Moore–Penrose inverse of Kronecker product function

$$\begin{aligned} m(\mathbb{f}^\dagger(A \otimes B)) &= \frac{\|\psi'(a, b)D_{a,b}\|_\infty}{\|\psi(a, b)\|_\infty} = \frac{\|[\phi_{(A,B)}L_B \quad \phi_{(A,B)}L_A] \begin{bmatrix} D_A & 0 \\ 0 & D_B \end{bmatrix}\|_\infty}{\|\text{vec}(\mathbb{f}^\dagger(A \otimes B))\|_\infty} \\ &= \frac{\|[\phi_{(A,B)}L_B \quad \phi_{(A,B)}L_A] \begin{bmatrix} D_A & 0 \\ 0 & D_B \end{bmatrix} \mathbf{e}\|_\infty}{\|\text{vec}(\mathbb{f}^\dagger(A \otimes B))\|_\infty} \\ &= \frac{\|[\phi_{(A,B)}L_B \quad \phi_{(A,B)}L_A] \begin{bmatrix} \text{vec}(|A|) \\ \text{vec}(|B|) \end{bmatrix}\|_\infty}{\|\text{vec}(\mathbb{f}^\dagger(A \otimes B))\|_\infty} \\ &= \frac{\|\phi_{(A,B)}L_B|\text{vec}(|A|) + \phi_{(A,B)}L_A|\text{vec}(|B|)\|_\infty}{\|\text{vec}(\mathbb{f}^\dagger(A \otimes B))\|_\infty}, \end{aligned}$$

and the componentwise condition number for  $f^\dagger(A \otimes B)$  is found to be

$$\begin{aligned} c(f^\dagger(A \otimes B)) &= \left\| D_{\psi(a,b)}^{-1} \psi'(a,b) D_{a,b} \right\|_\infty = \left\| D_{\text{vec}(f^\dagger(A \otimes B))}^{-1} [\phi_{(A,B)} L_B \phi_{(A,B)} L_A] \begin{bmatrix} D_A & 0 \\ 0 & D_B \end{bmatrix} \right\|_\infty \\ &= \left\| D_{\text{vec}(f^\dagger(A \otimes B))}^{-1} [\phi_{(A,B)} L_B \phi_{(A,B)} L_A] \begin{bmatrix} D_A & 0 \\ 0 & D_B \end{bmatrix} \mathbf{e} \right\|_\infty \\ &= \left\| \frac{[\phi_{(A,B)} L_B \phi_{(A,B)} L_A] \begin{bmatrix} \text{vec}(|A|) \\ \text{vec}(|B|) \end{bmatrix}}{\text{vec}(f^\dagger(A \otimes B))} \right\|_\infty \\ &= \left\| \frac{|\phi_{(A,B)} L_B| \text{vec}(|A|) + |\phi_{(A,B)} L_A| \text{vec}(|B|)}{\text{vec}(f^\dagger(A \otimes B))} \right\|_\infty, \end{aligned}$$

where  $\mathbf{e}$  is the vector of all ones with compatible dimension. □

The aforementioned theorem shows that the normwise condition number  $\kappa(f(A \otimes B))$  is proportional to the norms  $\|A\|_F$  and  $\|B\|_F$  and the reciprocal of  $\|f^\dagger(A \otimes B)\|_F$ , as expected.

Let us consider two matrix functions  $g(A) : \mathbb{R}^{m \times n} \mapsto \mathbb{R}^{e \times f}$  and  $h(B) : \mathbb{R}^{p \times q} \mapsto \mathbb{R}^{g \times h}$  and assume that  $f(A \otimes B) = g(A) \otimes h(B)$ . Then, it can be verified that  $f^\dagger(A \otimes B) = g^\dagger(A) \otimes h^\dagger(B)$ . Thus, we have

$$d(f^\dagger(A \otimes B)) = d(g^\dagger(A) \otimes h^\dagger(B)) = d(g^\dagger(A)) \otimes h^\dagger(B) + g^\dagger(A) \otimes d(h^\dagger(B)).$$

Vectorizing the aforementioned equation and using (2.2), we obtain

$$\begin{aligned} d(\text{vec}(g^\dagger(A) \otimes h^\dagger(B))) &= \text{vec}(d(g^\dagger(A)) \otimes h^\dagger(B) + g^\dagger(A) \otimes d(h^\dagger(B))) \\ &= (I_e \otimes K_{gf} \otimes I_h) \left\{ \text{vec}(d(g^\dagger(A))) \otimes \text{vec}(h^\dagger(B)) + \text{vec}(g^\dagger(A)) \otimes \text{vec}(d(h^\dagger(B))) \right\} \\ &= (I_e \otimes K_{gf} \otimes I_h) \left\{ (I_{fe} \otimes \text{vec}(h^\dagger(B))) \text{vec}(d(g^\dagger(A))) + (\text{vec}(g^\dagger(A)) \otimes I_{hg}) \text{vec}(d(h^\dagger(B))) \right\} \\ &= (I_e \otimes K_{gf} \otimes I_h) \left\{ (I_{fe} \otimes \text{vec}(h^\dagger(B))) \{ [g^{\dagger T}(A) g^\dagger(A) \otimes (I_f - g^\dagger(A) g(A)) \right. \\ &\quad + (I_e - g(A) g^\dagger(A)) \otimes g^\dagger(A) g^{\dagger T}(A) ] K_{ef} - (g^{\dagger T}(A) \otimes g^\dagger(A)) \} \text{vec}(d(g(A))) \\ &\quad + (\text{vec}(g^\dagger(A)) \otimes I_{hg}) \{ [h^{\dagger T}(B) h^\dagger(B) \otimes (I_h - h^\dagger(B) h(B)) \\ &\quad + (I_g - h(B) h^\dagger(B)) \otimes h^\dagger(B) h^{\dagger T}(B) ] K_{gh} - (h^{\dagger T}(B) \otimes h^\dagger(B)) \} \text{vec}(d(h(B))) \left. \right\}. \end{aligned} \tag{3.2}$$

Suppose that  $A$  and  $B$  are of full column rank. To present the condition numbers for the LS problems (1.1), we simply consider the function  $f(A \otimes B) = A \otimes B$  and obtain

$$g^\dagger(A) = A^\dagger, \quad h^\dagger(B) = B^\dagger \tag{3.3}$$

and

$$A^\dagger A = I_n, \quad B^\dagger B = I_q. \tag{3.4}$$

Analogous results can be obtained when  $A$  and  $B$  have full row rank.

Note that  $g(A) = A$  and  $h(B) = B$ ; thus,  $e = m$ ,  $f = n$ ,  $g = p$ , and  $h = q$ . As an immediate consequence of (3.2), by applying (3.3) and (3.4) we have the following corollary regarding the Fréchet derivative in the condition numbers for the Moore–Penrose inverse of a Kronecker product  $A \otimes B$ . Recall that  $\mathbb{R}_r^{m \times n}$  denotes the set of real  $m \times n$  matrices of rank  $r$ .

*Corollary 3.1*

Let  $A \in \mathbb{R}_n^{m \times n}$ ,  $B \in \mathbb{R}_q^{p \times q}$ ,  $\psi(a, b) = \text{vec}((A \otimes B)^\dagger)$ ,  $a = \text{vec}(A)$ , and  $b = \text{vec}(B)$ . Then  $\psi$  is continuous and Fréchet differentiable at all  $(a, b)$ . Moreover, it has the matrix expression

$$\psi'(a, b) = [Q_B M_A \quad P_A N_B],$$

where

$$M_A = -(A^{\dagger T} \otimes A^\dagger) + (I_m - AA^\dagger) \otimes (A^T A)^{-1} K_{mn}, \quad P_A = (I_m \otimes K_{pn} \otimes I_q)(\text{vec}(A^\dagger) \otimes I_{qp})$$

and

$$N_B = -(B^{\dagger T} \otimes B^\dagger) + (I_p - BB^\dagger) \otimes (B^T B)^{-1} K_{pq}, \quad Q_B = (I_m \otimes K_{pn} \otimes I_q)(I_{nm} \otimes \text{vec}(B^\dagger)).$$

From the aforementioned Fréchet derivative, we have the following corollary giving exact expressions of the condition numbers for the Moore–Penrose inverse of the Kronecker product  $A \otimes B$ , which will be useful for the symbolic computations in the absence of rounding errors.

*Corollary 3.2*

Let  $A \in \mathbb{R}_n^{m \times n}$  and  $B \in \mathbb{R}_q^{p \times q}$ . Respectively, the normwise and mixed condition numbers are of the following forms:

$$\begin{aligned} \kappa((A \otimes B)^\dagger) &= \frac{\|[Q_B M_A \quad P_A N_B]\|_2 \sqrt{\|A\|_F^2 + \|B\|_F^2}}{\|A^\dagger\|_F \|B^\dagger\|_F}, \\ m((A \otimes B)^\dagger) &= \frac{\|[Q_B M_A | \text{vec}(|A|) + |P_A N_B| \text{vec}(|B|)]\|_\infty}{\|\text{vec}(A^\dagger \otimes B^\dagger)\|_\infty}. \end{aligned}$$

Assuming there is no zero element in  $A^\dagger$  and  $B^\dagger$ , the componentwise condition number is expressed as

$$c((A \otimes B)^\dagger) = \left\| \frac{|Q_B M_A | \text{vec}(|A|) + |P_A N_B | \text{vec}(|B|)|}{\text{vec}(A^\dagger \otimes B^\dagger)} \right\|_\infty.$$

Corollary 3.2 gives explicit expressions for the condition numbers  $\kappa((A \otimes B)^\dagger)$ ,  $m((A \otimes B)^\dagger)$ , and  $c((A \otimes B)^\dagger)$ . When  $B = 1$ , they essentially reduce to the known  $\kappa(A^\dagger)$ ,  $m(A^\dagger)$ , and  $c(A^\dagger)$  in [1]. Although these expressions are exact using symbolic computation, they may not be easy to compute numerically because of the large size of the system involved and the commutation matrix  $K_{mn}$ . The following corollary presents simpler and practical upper bounds for the mixed and componentwise condition numbers.

*Corollary 3.3*

For the mixed and componentwise condition numbers in Corollary 3.2, the following simpler upper bounds hold:

$$\begin{aligned} m((A \otimes B)^\dagger) &\leq m((A \otimes B)^\dagger)^{\text{upper}} := \frac{\| |A^\dagger| \otimes (|B^\dagger| |B| |B^\dagger| + |(B^T B)^{-1}| |B|^T |I_p - BB^\dagger|) \|_{\max}}{\|A^\dagger \otimes B^\dagger\|_{\max}} \\ &\quad + \frac{\| (|A^\dagger| |A| |A^\dagger| + |(A^T A)^{-1}| |A|^T |I_m - AA^\dagger|) \otimes |B^\dagger| \|_{\max}}{\|A^\dagger \otimes B^\dagger\|_{\max}}, \\ c((A \otimes B)^\dagger) &\leq c((A \otimes B)^\dagger)^{\text{upper}} := \left\| \frac{|A^\dagger| \otimes (|B^\dagger| |B| |B^\dagger| + |(B^T B)^{-1}| |B|^T |I_p - BB^\dagger|)}{A^\dagger \otimes B^\dagger} \right\|_{\max} \\ &\quad + \left\| \frac{(|A^\dagger| |A| |A^\dagger| + |(A^T A)^{-1}| |A|^T |I_m - AA^\dagger|) \otimes |B^\dagger|}{A^\dagger \otimes B^\dagger} \right\|_{\max}. \end{aligned}$$

*Proof*

Using (2.1), (2.2), and (2.3), we have

$$\begin{aligned}
 |Q_B M_A \text{vec}(A)| &\leq |(I_m \otimes K_{pn} \otimes I_q)(I_{nm} \otimes \text{vec}(B^\dagger))| |[-(A^{\dagger T} \otimes A^\dagger) \\
 &\quad + (I_m - AA^\dagger) \otimes (A^T A)^{-1} K_{mn}] \text{vec}(A)| \\
 &\leq (I_m \otimes K_{pn} \otimes I_q)(I_{nm} \otimes \text{vec}(|B^\dagger|))(|A^{\dagger T}| \otimes |A^\dagger|) \\
 &\quad + |I_m - AA^\dagger| \otimes |(A^T A)^{-1} K_{mn}| \text{vec}(A)| \\
 &\leq (I_m \otimes K_{pn} \otimes I_q)(I_{nm} \otimes \text{vec}(|B^\dagger|)) \text{vec}(|A^\dagger| |A| |A^\dagger|) \\
 &\quad + |(A^T A)^{-1} |A|^T |I_m - AA^\dagger| \\
 &= (I_m \otimes K_{pn} \otimes I_q) \text{vec}(\text{vec}(|B^\dagger|) \text{vec}(|A^\dagger| |A| |A^\dagger| \\
 &\quad + |(A^T A)^{-1} |A|^T |I_m - AA^\dagger|)^T) \\
 &= (I_m \otimes K_{pn} \otimes I_q) \left[ \text{vec}(|A^\dagger| |A| |A^\dagger| \right. \\
 &\quad \left. + |(A^T A)^{-1} |A|^T |I_m - AA^\dagger| \otimes \text{vec}(|B^\dagger|) \right] \\
 &= \text{vec} \left( (|A^\dagger| |A| |A^\dagger| + |(A^T A)^{-1} |A|^T |I_m - AA^\dagger|) \otimes |B^\dagger| \right).
 \end{aligned}$$

Also, we can prove the following inequality:

$$|P_A N_B \text{vec}(|B|) \leq \text{vec} \left( |A^\dagger| \otimes (|B^\dagger| |B| |B^\dagger| + |(B^T B)^{-1} |B|^T |I_p - BB^\dagger|) \right).$$

Note that  $\|\text{vec}(A)\|_\infty = \|A\|_{\max}$  and the infinity norm is monotonic. The upper bounds can be obtained by applying the aforementioned two inequalities to the exact expressions of the condition numbers  $m((A \otimes B)^\dagger)$  and  $c((A \otimes B)^\dagger)$  in Corollary 3.2 and by using the matrix norm triangular inequality.  $\square$

#### 4. CONDITION NUMBERS FOR THE LINEAR LEAST SQUARES PROBLEM INVOLVING KRONECKER PRODUCTS

As we know, the LS problem

$$\min_{u \in \mathbb{R}^n} \|Au - b\|_2,$$

where  $A \in \mathbb{R}^{m \times n}$  is of full column rank and  $b \in \mathbb{R}^m$ , has a unique solution  $x = A^\dagger b$  [25, 32, 34–37]. Now that we have derived the condition numbers for the Moore–Penrose inverse  $(A \otimes B)^\dagger$  in Section 3, we are ready to study the condition numbers for the associated KPLS problem

$$\min_{v \in \mathbb{R}^{nq}} \|(A \otimes B)v - c\|_2, \tag{4.1}$$

where  $A \in \mathbb{R}_n^{m \times n}$ ,  $B \in \mathbb{R}_q^{p \times q}$ , and  $A \otimes B \in \mathbb{R}_{nq}^{mp \times nq}$ ,  $c \in \mathbb{R}^{mp}$ , (see e.g., [3–10] for details). It has a unique minimal 2-norm solution

$$x = (A \otimes B)^\dagger c = (A^\dagger \otimes B^\dagger)c = ((A^T A)^{-1} \otimes (B^T B)^{-1})(A^T \otimes B^T)c. \tag{4.2}$$

Our results generalize the perturbation analysis of the nonsingular linear equations  $(A \otimes B)x = d$  [38, 39] to the LS problem involving Kronecker products.

The perturbed system of (4.1) is

$$\min_{v \in \mathbb{R}^{nq}} \|[(A + \Delta A) \otimes (B + \Delta B)]v - (c + \Delta c)\|_2,$$

where  $\Delta A$ ,  $\Delta B$ , and  $\Delta c$  have the same dimensions as  $A$ ,  $B$ , and  $c$ , respectively.

Similar to Lemma 3.2, we first consider the Fréchet derivative.

*Lemma 4.1*

Let the mapping  $\varphi : \mathbb{R}^{mn} \times \mathbb{R}^{pq} \times \mathbb{R}^{mp} \mapsto \mathbb{R}^{nq}$  be defined by  $\varphi(a, b, c) = (A \otimes B)^\dagger c$ , where  $a = \text{vec}(A)$  and  $b = \text{vec}(B)$ . Assuming that  $A$  and  $B$  are of full column rank, then  $\varphi$  is continuous and Fréchet differentiable at all  $(\text{vec}(A), \text{vec}(B), c)$ . Moreover,

$$\varphi'(a, b, c) = [\mathcal{Q} \mathcal{P} A^\dagger \otimes B^\dagger],$$

where, denoting  $r = c - (A \otimes B)x$ , we obtain the following expressions for  $\mathcal{P}$  and  $\mathcal{Q}$ :

$$\begin{aligned} \mathcal{Q} &= (r^T \otimes ((A^T A)^{-1} \otimes (B^T B)^{-1}))(I_m \otimes K_{pn} \otimes I_q)(I_{mn} \otimes \text{vec}(B^T))K_{mn} \\ &\quad - (x^T \otimes (A^\dagger \otimes B^\dagger))(I_n \otimes K_{qm} \otimes I_p)(I_{mn} \otimes \text{vec}(B)), \\ \mathcal{P} &= (r^T \otimes ((A^T A)^{-1} \otimes (B^T B)^{-1}))(I_m \otimes K_{pn} \otimes I_q)(\text{vec}(A^T) \otimes I_{pq})K_{pq} \\ &\quad - (x^T \otimes (A^\dagger \otimes B^\dagger))(I_n \otimes K_{qm} \otimes I_p)(\text{vec}(A) \otimes I_{pq}). \end{aligned}$$

*Proof*

From Lemma 3.2, we know that  $A^\dagger \otimes B^\dagger$  is continuous and differentiable and so is  $\varphi(a, b, c) = (A \otimes B)^\dagger c$ . Differentiating both sides of (4.2), we have

$$\begin{aligned} dx &= d(((A^T A)^{-1} \otimes (B^T B)^{-1})(A^T \otimes B^T)c) \\ &= d(((A^T A)^{-1} \otimes (B^T B)^{-1})(A^T \otimes B^T)c) + d(((A^T A)^{-1} \otimes (B^T B)^{-1})d(A^T \otimes B^T)c) + (A^\dagger \otimes B^\dagger)dc \\ &= -((A^T A)^{-1} \otimes (B^T B)^{-1})d((A^T A) \otimes (B^T B))((A^T A)^{-1} \otimes (B^T B)^{-1})(A^T \otimes B^T)c \\ &\quad + ((A^T A)^{-1} \otimes (B^T B)^{-1})(dA^T \otimes B^T + A^T \otimes dB^T)c + (A^\dagger \otimes B^\dagger)dc \\ &= -((A^T A)^{-1} \otimes (B^T B)^{-1})((dA^T A) \otimes (B^T B) + (A^T dA) \otimes (B^T B) + (A^T A) \otimes (dB^T B) \\ &\quad + (A^T A) \otimes (B^T dB))x + ((A^T A)^{-1} \otimes (B^T B)^{-1})(dA^T \otimes B^T + A^T \otimes dB^T)c + (A^\dagger \otimes B^\dagger)dc \\ &= -(A^\dagger \otimes B^\dagger)(dA \otimes B + A \otimes dB)x - ((A^T A)^{-1} \otimes (B^T B)^{-1})(dA^T \otimes B^T + A^T \otimes dB^T)(A \otimes B)x \\ &\quad + ((A^T A)^{-1} \otimes (B^T B)^{-1})(dA^T \otimes B^T + A^T \otimes dB^T)c + (A^\dagger \otimes B^\dagger)dc \\ &= -(A^\dagger \otimes B^\dagger)(dA \otimes B + A \otimes dB)x + ((A^T A)^{-1} \otimes (B^T B)^{-1})(dA^T \otimes B^T + A^T \otimes dB^T)r \\ &\quad + (A^\dagger \otimes B^\dagger)dc. \end{aligned} \tag{4.3}$$

We then vectorize both sides of the aforementioned equation and obtain

$$\begin{aligned} dx &= \text{vec}(dx) = \text{vec}(-(A^\dagger \otimes B^\dagger)(dA \otimes B + A \otimes dB)x) \\ &\quad + \text{vec}(((A^T A)^{-1} \otimes (B^T B)^{-1})(dA^T \otimes B^T + A^T \otimes dB^T)r) + \text{vec}((A^\dagger \otimes B^\dagger)dc) \\ &= -(x^T \otimes (A^\dagger \otimes B^\dagger))(I_n \otimes K_{qm} \otimes I_p)[\text{vec}(dA) \otimes \text{vec}(B) + \text{vec}(A) \otimes \text{vec}(dB)] \\ &\quad + (r^T \otimes ((A^T A)^{-1} \otimes (B^T B)^{-1}))(I_m \otimes K_{pn} \otimes I_q)[\text{vec}(dA^T) \otimes \text{vec}(B^T) \\ &\quad + \text{vec}(A^T) \otimes \text{vec}(dB^T)] + (A^\dagger \otimes B^\dagger)dc \\ &= -(x^T \otimes (A^\dagger \otimes B^\dagger))(I_n \otimes K_{qm} \otimes I_p)[(I_{mn} \otimes \text{vec}(B))\text{vec}(dA) + (\text{vec}(A) \otimes I_{pq})\text{vec}(dB)] \\ &\quad + (r^T \otimes ((A^T A)^{-1} \otimes (B^T B)^{-1}))(I_m \otimes K_{pn} \otimes I_q)[(I_{mn} \otimes \text{vec}(B^T))K_{mn}\text{vec}(dA) \\ &\quad + (\text{vec}(A^T) \otimes I_{pq})K_{pq}\text{vec}(dB)] + (A^\dagger \otimes B^\dagger)dc. \end{aligned}$$

That is,

$$dx = [\mathcal{Q} \mathcal{P} A^\dagger \otimes B^\dagger][da^T \ db^T \ dc^T]^T.$$

So, the Fréchet derivative is  $\varphi'(a, b, c) = [\mathcal{Q} \mathcal{P} A^\dagger \otimes B^\dagger]$ . □

From the aforementioned Fréchet derivative, we have the following exact expressions of the condition numbers for the KPLS problem and their upper bounds, which can be used to efficiently estimate the condition numbers.

*Corollary 4.1*

Using the notations in Lemma 4.1, we have the following normwise and mixed condition numbers and their upper bounds:

$$\begin{aligned}
 \kappa^{ls}(A \otimes B, c) &:= \lim_{\epsilon \rightarrow 0} \sup \frac{\|\Delta x\|_2}{\epsilon \|x\|_2} \\
 &\quad \frac{\sqrt{\|\Delta A\|_F^2 + \|\Delta B\|_F^2 + \|\Delta c\|_2^2}}{\sqrt{\|A\|_F^2 + \|B\|_F^2 + \|c\|_2^2}} \\
 &\leq \epsilon \sqrt{\|A\|_F^2 + \|B\|_F^2 + \|c\|_2^2} \\
 &= \frac{\|[\mathcal{Q} \ \mathcal{P} \ A^\dagger \otimes B^\dagger]\|_2 \sqrt{\|A\|_F^2 + \|B\|_F^2 + \|c\|_2^2}}{\|x\|_2} \\
 m^{ls}(A \otimes B, c) &:= \lim_{\epsilon \rightarrow 0} \sup \frac{1}{\epsilon} \frac{\|\Delta x\|_\infty}{\|x\|_\infty} \\
 &\quad \begin{aligned} &|\Delta A| \leq \epsilon |A| \\ &|\Delta B| \leq \epsilon |B| \\ &|\Delta c| \leq \epsilon |c| \end{aligned} \\
 &= \frac{\|[\mathcal{Q}|\text{vec}(|A|) + |\mathcal{P}|\text{vec}(|B|) + (|A^\dagger| \otimes |B^\dagger|)|c|]\|_\infty}{\|x\|_\infty} \\
 &\leq \frac{2 \| |A^\dagger| \otimes |B^\dagger| (|A| \otimes |B|) |x| \|_\infty}{\|x\|_\infty} \\
 &\quad + \frac{2 \| |(A^T A)^{-1} \otimes (B^T B)^{-1} (|A|^T \otimes |B|^T) |r| \|_\infty}{\|x\|_\infty} \\
 &\quad + \frac{\| |A^\dagger| \otimes |B^\dagger| |c| \|_\infty}{\|x\|_\infty} \\
 &=: m^{ls}(A \otimes B, c)^{\text{upper}}.
 \end{aligned}$$

If there is no zero element in  $x$ , then the componentwise condition number is

$$\begin{aligned}
 c^{ls}(A \otimes B, c) &:= \lim_{\epsilon \rightarrow 0} \sup \frac{1}{\epsilon} \left\| \frac{\Delta x}{x} \right\|_\infty \\
 &\quad \begin{aligned} &|\Delta A| \leq \epsilon |A| \\ &|\Delta B| \leq \epsilon |B| \\ &|\Delta c| \leq \epsilon |c| \end{aligned} \\
 &= \left\| \frac{[\mathcal{Q}|\text{vec}(|A|) + |\mathcal{P}|\text{vec}(|B|) + (|A^\dagger| \otimes |B^\dagger|)|c|]}{x} \right\|_\infty \\
 &\leq 2 \left\| |D_x^{-1}| |A^\dagger| \otimes |B^\dagger| (|A| \otimes |B|) |x| \right\|_\infty \\
 &\quad + 2 \left\| |D_x^{-1}| |(A^T A)^{-1} \otimes (B^T B)^{-1} (|A|^T \otimes |B|^T) |r| \right\|_\infty \\
 &\quad + \left\| |D_x^{-1}| |A^\dagger| \otimes |B^\dagger| |c| \right\|_\infty \\
 &=: c^{ls}(A \otimes B, c)^{\text{upper}}.
 \end{aligned}$$

*Proof*

Combining Lemmas 2.1 and 4.1, we obtain the exact expressions for the normwise, mixed, and componentwise condition numbers. Using (2.1), (2.2), and (2.3), we have

$$\begin{aligned}
|\mathcal{Q}||\text{vec}(A)| &\leq |(r^T \otimes ((A^T A)^{-1} \otimes (B^T B)^{-1}))(I_m \otimes K_{pn} \otimes I_q)(I_{mn} \otimes \text{vec}(B^T))K_{mn}| |\text{vec}(A)| \\
&\quad + |(x^T \otimes (A^\dagger \otimes B^\dagger))(I_n \otimes K_{qm} \otimes I_p)(I_{mn} \otimes \text{vec}(B))| |\text{vec}(A)| \\
&\leq [ |r^T| \otimes |(A^T A)^{-1} \otimes (B^T B)^{-1}| (I_m \otimes K_{pn} \otimes I_q)(I_{mn} \otimes \text{vec}(|B^T|)) ] \text{vec}(|A^T|) \\
&\quad + [ |x^T| \otimes |A^\dagger \otimes B^\dagger| (I_n \otimes K_{qm} \otimes I_p)(I_{mn} \otimes \text{vec}(|B|)) ] \text{vec}(|A|) \\
&= (|r^T| \otimes |(A^T A)^{-1} \otimes (B^T B)^{-1}|)(I_m \otimes K_{pn} \otimes I_q) \text{vec}(\text{vec}(|B^T|)\text{vec}(|A^T|)^T) \\
&\quad + (|x^T| \otimes |A^\dagger \otimes B^\dagger|)(I_n \otimes K_{qm} \otimes I_p) \text{vec}(\text{vec}(|B|)\text{vec}(|A|)^T) \\
&= (|r^T| \otimes |(A^T A)^{-1} \otimes (B^T B)^{-1}|) \text{vec}(|A^T| \otimes |B^T|) + (|x^T| \otimes |A^\dagger \otimes B^\dagger|) \text{vec}(|A| \otimes |B|) \\
&= |A^\dagger \otimes B^\dagger|(|A| \otimes |B|)|x| + |(A^T A)^{-1} \otimes (B^T B)^{-1}|(|A|^T \otimes |B|^T)|r|.
\end{aligned}$$

Similarly, we can deduce that

$$|\mathcal{P}|\text{vec}(|B|) \leq |A^\dagger \otimes B^\dagger|(|A| \otimes |B|)|x| + |(A^T A)^{-1} \otimes (B^T B)^{-1}|(|A|^T \otimes |B|^T)|r|.$$

Because of the monotonicity property of the infinity norm, the upper bounds  $m^{ls}(A \otimes B, c)^{\text{upper}}$  and  $c^{ls}(A \otimes B, c)^{\text{upper}}$  can be obtained by applying the aforementioned two inequalities to the exact expressions of  $m^{ls}(A \otimes B, c)$  and  $c^{ls}(A \otimes B, c)$  and by using the matrix norm triangular inequality.  $\square$

*Remark 4.1*

From the small example in Section 1, the derived upper bounds  $m^{ls}(A \otimes B, c)^{\text{upper}}$  and  $c^{ls}(A \otimes B, c)^{\text{upper}}$  are achievable.

Finally, we present an approach to compute the upper bounds, in particular the second term in  $m^{ls}(A \otimes B, c)^{\text{upper}}$  or  $c^{ls}(A \otimes B, c)^{\text{upper}}$ , involving  $\| |(A^T A)^{-1} \otimes (B^T B)^{-1}|(|A|^T \otimes |B|^T)|r| \|_\infty$ . Let  $z = (|A|^T \otimes |B|^T)|r|$  and  $\mathbf{e}$  be the vector of all ones, then we have

$$\begin{aligned}
&\| |(A^T A)^{-1} \otimes (B^T B)^{-1}|(|A|^T \otimes |B|^T)|r| \|_\infty \\
&= \| |(A^T A)^{-1} \otimes (B^T B)^{-1}|z \|_\infty \\
&= \| |(A^T A)^{-1} \otimes (B^T B)^{-1}|D_z \mathbf{e} \|_\infty \\
&= \| |(A^T A)^{-1} \otimes (B^T B)^{-1}|D_z \|_\infty \\
&= \| ((A^T A)^{-1} \otimes (B^T B)^{-1})D_z \|_\infty.
\end{aligned}$$

We can apply LAPACK's condition estimator SLACON, based on the Hager–Higham's algorithm [26, 27], to efficiently estimate the aforementioned matrix  $\infty$ -norm. The estimator involves solutions of some linear systems with the coefficient matrix  $(A^T A) \otimes (B^T B)$ . If the Cholesky factors  $R_1$  and  $R_2$  of  $A^T A$  and  $B^T B$ , respectively, are available, then

$$(A^T A) \otimes (B^T B) = (R_1^T \otimes R_2^T)(R_1 \otimes R_2).$$

Thus, the upper bounds can be efficiently estimated by exploiting the triangular structure of  $R_1$  and  $R_2$ .

## 5. NUMERICAL EXAMPLE

In this section, we present test results with our computable upper bounds for the condition numbers of the bivariate least squares tensor product problem. We compare the computable bounds of the

Table I. Kronecker products linear least squares problems.

$m$	$n$	$\kappa^{ls}(V_1 \otimes V_2, f)$	$m^{ls}(V_1 \otimes V_2, f)$	$c^{ls}(V_1 \otimes V_2, f)$
			$m^{ls}(V_1 \otimes V_2, f)^{\text{upper}}$	$c^{ls}(V_1 \otimes V_2, f)^{\text{upper}}$
4	3	1.4013e+002	7.7233e+001	2.7361e+002
			7.3825e+003	2.9440e+004
8	5	1.2264e+004	3.8985e+003	2.6674e+005
			5.1227e+006	3.6968e+008
12	6	3.6263e+004	5.1042e+003	2.4146e+004
			7.0622e+007	2.0248e+008

condition numbers with their corresponding exact values. All computations were carried out using MATLAB (MathWorks, Natick, MA, USA) 7.0 with precision  $\epsilon \approx 2.2 \times 10^{-16}$ . In our experiments, when computing the upper bounds, the computed solution was used as an estimation of the exact solution  $x$ .

*Example 5.1* ([13])

For  $i = 1, 2$ , consider intervals  $I_i = [a_i, b_i]$ , and let  $E_i$  be finite-dimensional subspaces of all functions from  $I_i$  to  $\mathbb{R}$ . Let  $n_i = \dim(E_i)$ ,  $x_1, x_2, \dots, x_{m_1} \in [a_1, b_1]$ ,  $y_1, y_2, \dots, y_{m_2} \in [a_2, b_2]$  pairwise different in each case. The bivariate least squares tensor product problem is, for a given  $f \in \mathbb{R}^{m_1 \cdot m_2}$ , to find a tensor product function  $g \in E_1 \otimes E_2$  minimizing

$$\sum_{p=1}^{m_1} \sum_{q=1}^{m_2} [g(x_p, y_q) - f_{(p-1)m_2+q}]^2. \tag{5.1}$$

The aforementioned minimization problem can be rewritten as  $\min_{x \in \mathbb{R}^{n_1 \cdot n_2}} \|(V_1 \otimes V_2)x - f\|_2$ .

In the numerical experiments, we choose  $I_1 = I_2 = [0, 1]$ ,  $x_i = y_i = \frac{i}{m}$ ,  $i = 1, 2, \dots, m$ , and  $E_1 = E_2$  as the space of all polynomials of degree  $\leq (n - 1)$ . Let  $f$  be a random vector with proper size. The comparison is given in Table I where each two rows correspond to an array of parameters  $(m, n)$ .

Table I shows that the upper bounds are around three orders of magnitude larger than their exact counterparts.

We also test our computable bounds for other problems with different settings and obtain results similar to those in the aforementioned example. That is, the computable upper bounds are around three orders of magnitude larger than their exact counterparts.

## 6. CONCLUSION REMARKS

This paper presents three types of condition numbers of the Moore–Penrose inverse of a matrix with elements generated by a single Kronecker matrix product. From them, the condition numbers of the Moore–Penrose inverse of a Kronecker product and the associated least squares problem are derived. Computable upper bounds for the condition numbers are also provided. It is of interest to extend our results to the corresponding rank-deficient and weighted KPLS problems [25, 40] and to develop extra-precise iterative refinement [34, 41] for such problems.

### ACKNOWLEDGEMENTS

The authors would like to thank Dr. Maya Neytcheva and Prof. Lars Eldén for their useful comments and Mr. Hai-ming Zhang for providing us with numerical examples. We wish to thank the two anonymous referees for their constructive and detailed comments that enabled us to substantially improve our presentation.

H. Diao is supported by the National Natural Science Foundation of China under grant 11001045 and by the Youth Foundation of Tianyuan Mathematics, National Natural Science Foundation of China under grant 10926107, Specialized Research Fund for the Doctoral Program of Higher Education of Ministry of Education, China under grant 20090043120008, and Training Fund of NENU Scientific Innovation Project of Northeast Normal University under grant NENU-STC08009.

W. Wang is supported by the National Natural Science Foundation of China under grant 10971204, China Scholarship Council, the Shandong Province Natural Science Foundation (Y2008A07), and Fundamental Research Funds for the Central Universities (201013048). Partial work was finished when he visited University of Texas at Arlington and Shanghai Key Laboratory of Contemporary Applied Mathematics.

Y. Wei is supported by the National Natural Science Foundation of China under grant 10871051, Doctoral Program of the Ministry of Education under grant 20090071110003, 973 Program Project under grant 2010CB327900, and Shanghai Education Committee (Dawn Project).

Partial work was finished when S. Qiao visited Shanghai Key Laboratory of Contemporary Applied Mathematics.

#### REFERENCES

1. Cucker F, Diao H, Wei Y. On mixed and componentwise condition numbers for Moore-Penrose inverse and linear least squares problems. *Mathematics of Computation* 2007; **76**:947–963.
2. Xu W, Wei Y, Qiao S. Condition numbers for structured least squares problems. *BIT* 2006; **46**:203–225.
3. Barrlund A. Efficient solution of constrained least squares problems with Kronecker product structure. *SIAM Journal on Matrix Analysis and Applications* 1998; **19**:154–160.
4. Kamm J, Nagy J. Optimal Kronecker product approximation of block Toeplitz matrices. *SIAM Journal on Matrix Analysis and Applications* 2000; **22**:155–172.
5. Kang W, Xiang H. Level-2 condition numbers for least-squares solution of Kronecker product linear systems. *International Journal of Computer Mathematics* 2008; **85**:827–841.
6. Langville AN, Stewart WJ. A Kronecker product approximate preconditioner for SANs. *Numerical Linear Algebra with Applications* 2004; **11**:723–752.
7. Langville AN, Stewart WJ. The Kronecker product and stochastic automata networks. *Journal of Computational and Applied Mathematics* 2004; **167**:429–447.
8. Nagy J, Ng MK, Perrone L. Kronecker product approximations for image restoration with reflexive boundary conditions. *SIAM Journal on Matrix Analysis and Applications* 2003; **25**:829–841.
9. Van Loan CF. The ubiquitous Kronecker product. *Journal of Computational and Applied Mathematics* 2000; **123**:85–100.
10. Zha H. Comments on: large least squares problems involving Kronecker products. *SIAM Journal on Matrix Analysis and Applications* 1995; **16**:1172.
11. Oh S, Kwon S, Yun JH. A method for structured linear total least norm on blind deconvolution problem. *Journal of Applied Mathematics and Computing* 2005; **19**:151–164.
12. Pisinger G, Zimmermann A. Bivariate least squares approximation with linear constraints. *BIT* 2007; **47**:427–439.
13. Pisinger G, Zimmermann A. Linear least squares problems with data over incomplete grids. *BIT* 2007; **47**:809–824.
14. Anderson E, Bai Z, Bischof C, Blackford S, Demmel J, Dongarra J, Du Croz J, Greenbaum A, Hammarling S, McKenney A, Sorensen D. (Edits), *LAPACK Users' Guide*, Third Edition. SIAM, Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1999.
15. Gohberg I, Koltracht I. Mixed, componentwise, and structured condition numbers. *SIAM Journal on Matrix Analysis and Applications* 1993; **14**:688–704.
16. Skeel RD. Scaling for numerical stability in Gaussian elimination. *Journal of the Association for Computing Machinery* 1979; **26**:817–526.
17. Rohn J. New condition numbers for matrices and linear systems. *Computing* 1989; **41**:167–169.
18. Arioli M, Duff IS, de Rijk PPM. An augmented system approach to sparse least-squares problems. *Numerische Mathematik* 1989; **55**:667–684.
19. Björck Å. Component-wise perturbation analysis and error bounds for linear least squares solutions. *BIT* 1991; **31**:238–244.
20. Rump SM. Structured perturbations part II: componentwise distances. *SIAM Journal on Matrix Analysis and Applications* 2003; **25**:31–56.
21. Cucker F, Diao H. Mixed and componentwise condition numbers for rectangular structured matrices. *Calcolo* 2007; **44**:89–115.
22. Arioli M, Baboulin M, Gratton S. A partial condition number for linear least squares problems. *SIAM Journal on Matrix Analysis and Applications* 2007; **29**:413–433.
23. Baboulin M, Gratton S. Using dual techniques to derive componentwise and mixed condition numbers for a linear function of a linear least squares solution. *BIT* 2009; **49**:3–19.
24. Wang D, Zhi L (eds). *Symbolic-Numeric Computation*, Trends in Mathematics. Birkhäuser Verlag: Basel, 2007.
25. Björck Å. *Numerical Methods for Least Squares Problems*, Society for Industrial and Applied Mathematics. SIAM: Philadelphia, PA, 1996.
26. Hager WW. Condition estimates. *SIAM Journal on Scientific and Statistical Computing* 1984; **5**:311–316.

27. Higham NJ. FORTRAN codes for estimating the one-norm of a real or complex matrix, with applications to condition estimation (Algorithm 674). *ACM Transactions on Mathematical Software* 1988; **14**:381–396.
28. Rice JR. A theory of condition. *SIAM Journal on Numerical Analysis* 1966; **3**:217–232.
29. Graham A. *Kronecker Products and Matrix Calculus with Applications*. Ellis Horwood Ltd: New York, NY, USA, 1981.
30. Schott JR. *Matrix Analysis for Statistics*, 2nd Edition. John Wiley Sons, Inc: Hoboken, NJ, USA, 2005.
31. Ben-Israel A, Greville TNE. *Generalized Inverses: Theory and Applications*, 2nd Edition. Springer Verlag: New York, 2003.
32. Golub GH, Van Loan CF. *Matrix Computations*, 3rd Edition. Johns Hopkins University: Baltimore, 1996.
33. Magnus JR, Neudecker H. *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Third Edition. John Wiley Sons Inc.: Hoboken, NJ, USA, 2007.
34. Demmel J, Hida Y, Li XS, Riedy E. Extra-precise iterative refinement for overdetermined least squares problems. *ACM Transactions on Mathematical Software* 2009; **35**(4).
35. Jin X. *Developments and Applications of Block Toeplitz Iterative Solvers*. Kluwer Academic Publishers Group, Dordrecht; Science Press: Beijing, 2002.
36. Higham NJ. *Accuracy and Stability of Numerical Algorithms*, 2nd edition. SIAM: Philadelphia, PA, 2002.
37. Wedin P-Å. Perturbation theory for pseudo-inverses. *BIT* 1973; **13**:217–232.
38. Kressner D, Tobler C. Krylov subspace methods for linear systems with Tensor product structure. *SIAM Journal on Matrix Analysis and Applications* 2010; **31**:1688–1714.
39. Xiang H, Diao H, Wei Y. On perturbation bounds of Kronecker product linear systems and their level-2 condition numbers. *Journal of Computational and Applied Mathematics* 2005; **183**:210–231.
40. Wei M. *Supremum and Stability of Weighted Pseudoinverses and Weighted Least Squares Problems, Analysis and Computations*. Nova Science Publishers: Huntington, NY, 2001.
41. Baboulin M, Dongarra J, Gratton S, Langou J. Computing the conditioning of the components of a linear least-squares solution. *Numerical Linear Algebra with Applications* 2009; **16**:517–533.