Operators and Matrices
Volume 15, Number 2 (2021), 645–657 doi:10.7153/oam-2021-15-43

FROBENIUS NORM INEQUALITIES OF COMMUTATORS BASED ON DIFFERENT PRODUCTS

WEI-HUI LIU, ZE-JIA XIE AND XIAO-QING JIN*

(*Communicated by F. Kittaneh*)

Abstract. The difference $AB-BA$ of two matrices A and B is called the commutator (or Lie product). In this paper, we are concerned with inequalities for the Frobenius norm of commutators based on other products, including the Kronecker product, the Khatri-Rao product, the contracted product, and the T-product. We also study the characterization of their corresponding maximal pairs.

1. Introduction

The commutator plays an important role in the areas of Lie group, Lie algebra, perturbation analysis, operator theory, and matrix manifold computation [1, 4, 5, 25, 26, 38]. In 2005, Böttcher and Wenzel [6] proposed a conjecture that the upper bound for the Frobenius norm of the commutator of any $A,B \in \mathbb{R}^{n\times n}$ is given by

$$\|AB-BA\|_F \leqslant \sqrt{2}\|A\|_F\|B\|_F. \tag{1.1}$$

Note that $\sqrt{2}$ is the best possible coefficient, because the equality in (1.1) holds when we take

$$A=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}, \quad B=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}.$$

Besides, Böttcher and Wenzel [6] also proved the conjecture for the case of 2×2 matrices. In 2007, László [27] showed that the inequality holds for the case of 3×3 matrices. The conjecture was first proved for any pair of $n\times n$ real matrices in 2008 by Vong and Jin [36]. Later, Lu [29] gave a different proof independently and the result is included in [30]. Böttcher and Wenzel [7] extended the result to general $n\times n$ complex matrices. Some other alternative proofs of this conjecture can be found in [2, 31].

Mathematics subject classification (2020): 15A52, 15A45, 60H25, 65F35.
Keywords and phrases: Commutator, Böttcher-Wenzel conjecture, Frobenius norm, Kronecker product, Khatri-Rao product, contracted product, T-product, maximal pair.
The research is supported by research grants MYRG2019-00042-FST from University of Macau, 0014/2019/A from FDCT of Macao, and National Natural Science Foundation of China (No. 11801074).
* Corresponding author.

After the affirmation of the Böttcher-Wenzel conjecture, there were several subsequent problems considered. One is the maximal pairs of the inequality. By maximal pairs, one addresses those two nonzero matrices A and B such that

$$\|AB - BA\|_F = \sqrt{2}\|A\|_F\|B\|_F,$$

i.e., (1.1) holds in (non-trivial) equality.

Böttcher and Wenzel [7] gave some necessary conditions for maximal pairs and obtained maximal pairs for 2×2 matrices, rank one matrices, and normal matrices. Later, Cheng, Vong, and Wenzel [15] gave a complete characterization of maximal pairs. We remark that the proof in [15] was deduced by heavy calculations. Cheng et al. [9] presented an alternative proof of the characterization of maximal pairs again according to a proof in [2]. For convenience of the study in subsequent sections, we similarly call a pair of nonzero matrices (or tensors) satisfying the respective equality a maximal pair for the corresponding inequality. For more research on norm inequalities of the commutator, we refer to [8, 10, 11, 12, 13, 14, 18, 19, 21, 32, 37].

In this paper, we study the inequality (1.1) based on other products than the usual matrix product and propose the characterization of their maximal pairs.

2. Commutators based on other matrix products

In this section, we study the norm inequalities of commutators based on the Kronecker product and the Khatri-Rao product, and we investigate how the different products relate.

2.1. Kronecker product

Given two matrices $A=[a_{ij}]\in\mathbb{C}^{m\times n}$ and $B=[b_{ij}]\in\mathbb{C}^{p\times q}$, the Kronecker product of A and B is defined by

$$A\otimes B := \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \in \mathbb{C}^{mp\times nq}.$$

The Frobenius norm inequality of commutator based on the Kronecker product has been studied by Böttcher and Wenzel [7] for square matrices. Their result is easily extended to rectangular matrices.

THEOREM 2.1. *Let $A\in\mathbb{C}^{m\times n}$ and $B\in\mathbb{C}^{m\times n}$. We have*

$$\|A\otimes B - B\otimes A\|_F \leqslant \sqrt{2}\|A\|_F\|B\|_F. \tag{2.1}$$

Furthermore, when A,B are nonzero, (A,B) is a maximal pair for (2.1) if and only if $\operatorname{tr}(A^*B)=0$.

The proof of the theorem is similar to that in [7] and we therefore omit it.

EXAMPLE 2.2. We present a kind of maximal pair for (2.1). Let $A,B \in \mathbb{C}^{3\times 2}$, where

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$

with $b_{11}+b_{22}=0$. It is easy to see that $\operatorname{tr}(A^*B) = b_{11}+b_{22} = 0$ and $|b_{11}| = |b_{22}|$. By some calculations, we find that

$$\begin{aligned} \|A\otimes B - B\otimes A\|_F^2 &= 2|b_{11}-b_{22}|^2 + 4\sum_{i\neq j}|b_{ij}|^2 \\ &= 8|b_{11}|^2 + 4\sum_{i\neq j}|b_{ij}|^2 \\ &= 4\sum_{i,j}|b_{ij}|^2 = 2\|A\|_F^2\|B\|_F^2. \end{aligned}$$

2.2. Khatri-Rao product

Next we discuss the commutator based on another product of matrices known as the Khatri-Rao product [22]. It was first presented by Khatri and Rao in 1968, which was used to solve some functional equations of the characterization of probability distribution [22] and the estimation of heteroscedastic variances [35]. For matrices

$$A = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n] \in \mathbb{C}^{m\times n}, \quad B = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n] \in \mathbb{C}^{p\times n}$$

with the same number of columns n, the Khatri-Rao product $A\odot B$ is defined to be the partitioned matrix

$$A\odot B := [\mathbf{a}_1\otimes\mathbf{b}_1, \mathbf{a}_2\otimes\mathbf{b}_2, \ldots, \mathbf{a}_n\otimes\mathbf{b}_n] \in \mathbb{C}^{mp\times n}.$$

THEOREM 2.3. *Let $A\in\mathbb{C}^{m\times n}$, $B\in\mathbb{C}^{m\times n}$. We have*

$$\|A\odot B - B\odot A\|_F \leqslant \sqrt{2}\|A\|_F\|B\|_F. \tag{2.2}$$

Proof. By the previous definition, we have

$$A\odot B - B\odot A = [\mathbf{a}_1\otimes\mathbf{b}_1 - \mathbf{b}_1\otimes\mathbf{a}_1, \ldots, \mathbf{a}_n\otimes\mathbf{b}_n - \mathbf{b}_n\otimes\mathbf{a}_n] \in \mathbb{C}^{m^2\times n}$$

and

$$\|A\odot B - B\odot A\|_F^2 = \|\mathbf{a}_1\otimes\mathbf{b}_1 - \mathbf{b}_1\otimes\mathbf{a}_1\|_2^2 + \cdots + \|\mathbf{a}_n\otimes\mathbf{b}_n - \mathbf{b}_n\otimes\mathbf{a}_n\|_2^2. \tag{2.3}$$

From the inequality (2.1) for the Kronecker product, one has

$$\|\mathbf{a}_j\otimes\mathbf{b}_j - \mathbf{b}_j\otimes\mathbf{a}_j\|_2^2 \leqslant 2\|\mathbf{a}_j\|_2^2\|\mathbf{b}_j\|_2^2,$$

for $1 \leqslant j \leqslant n$. Thus, it follows from (2.3) that

$$\begin{aligned}\|A \odot B - B \odot A\|_F^2 &\leqslant 2\,(\|\mathbf{a}_1\|_2^2\|\mathbf{b}_1\|_2^2 + \cdots + \|\mathbf{a}_n\|_2^2\|\mathbf{b}_n\|_2^2)\\ &\leqslant 2\,(\|\mathbf{a}_1\|_2^2 + \cdots + \|\mathbf{a}_n\|_2^2)(\|\mathbf{b}_1\|_2^2 + \cdots + \|\mathbf{b}_n\|_2^2)\\ &= 2\,\|A\|_F^2\|B\|_F^2.\end{aligned}$$

Hence, (2.2) is immediate. □

The characterization of maximal pairs for (2.2) is a little more restrictive compared to the one of (2.1), taking into account the additional structures.

COROLLARY 2.4. *Let* $A = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n], B = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n] \in \mathbb{C}^{m\times n}$ *be nonzero. Then* (A,B) *is a maximal pair for* (2.2) *if and only if there exists an integer* i *with* $1 \leqslant i \leqslant n$ *such that*

(1) $\mathbf{a}_i^*\mathbf{b}_i = 0$;

(2) $\mathbf{a}_j = \mathbf{b}_j = \mathbf{0}$, *for any* $j \neq i$, $1 \leqslant j \leqslant n$.

Proof. Based on the proof of above theorem, we find that (A,B) is a maximal pair for (2.2) if and only if

(i) $\mathbf{a}_k^*\mathbf{b}_k = \mathrm{tr}\,(\mathbf{a}_k^*\mathbf{b}_k) = 0$, for every $1 \leqslant k \leqslant n$;

(ii) $\|\mathbf{a}_k\|_2^2\|\mathbf{b}_j\|_2^2 = 0$, for any $k \neq j$, $1 \leqslant k, j \leqslant n$.

Since A and B are nonzero, suppose $\mathbf{a}_i \neq \mathbf{0}$ for a fixed i. According to condition (ii), we get $\|\mathbf{b}_j\|_2^2 = 0$, i.e., $\mathbf{b}_j = \mathbf{0}$, for any $j \neq i$. Since $B \neq \mathbf{0}$, we have $\mathbf{b}_i \neq \mathbf{0}$. By using condition (ii) again, we obtain $\mathbf{a}_j = \mathbf{0}$, for any $j \neq i$. Hence only $\mathbf{a}_i$ and $\mathbf{b}_i$ are nonzero, and the other columns are zero. Therefore, (1) & (2) is equivalent to (i) & (ii). □

EXAMPLE 2.5. We present a simple example of a maximal pair for (2.2). Let $A, B \in \mathbb{R}^{3\times 2}$, where

$$A = \begin{bmatrix} 0 & 1\\ 0 & 0\\ 0 & 1\end{bmatrix}, \quad B = \begin{bmatrix} 0 & b_{12}\\ 0 & b_{22}\\ 0 & b_{32}\end{bmatrix}$$

with $b_{12} + b_{32} = 0$. Thus, $\mathbf{a}_1 = \mathbf{b}_1 = \mathbf{0}$ and $\mathbf{a}_2^*\mathbf{b}_2 = b_{12} + b_{32} = 0$. By some calculations, we have

$$\begin{aligned}\|A \odot B - B \odot A\|_F^2 &= 2|b_{12} - b_{32}|^2 + 4|b_{22}|^2\\ &= 8|b_{12}|^2 + 4|b_{22}|^2\\ &= 4(|b_{12}|^2 + |b_{22}|^2 + |b_{32}|^2)\\ &= 2\|A\|_F^2\|B\|_F^2.\end{aligned}$$

2.3. Relationship of commutators based on different products

The following theorem introduces a chain of inequalities of the commutators based on the matrix product and the Kronecker product.

THEOREM 2.6. [7] *If* $A,B \in \mathbb{C}^{n\times n}$, *then*

$$\|AB - BA\|_F \leqslant \|A\otimes B - B\otimes A\|_F \leqslant \sqrt{2}\|A\|_F\|B\|_F.$$

A similar relation can be found with the Khatri-Rao product.

THEOREM 2.7. *Let* $A,B \in \mathbb{C}^{m\times n}$, *then we have*

$$\|A\odot B - B\odot A\|_F \leqslant \|A\otimes B - B\otimes A\|_F.$$

Proof. It is easy to see that

$$A\otimes B = \left[\mathbf{a}_1\otimes B \mid \mathbf{a}_2\otimes B \mid \ldots \mid \mathbf{a}_n\otimes B\right],$$

where A is partitioned as $A = [\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_n]$. Thus, every column of $A\odot B$ is a column of $A\otimes B$. This is also preserved under subtracting two such objects, completing the proof. □

However, the commutators based on the matrix product and the Khatri-Rao product cannot be compared directly. The following example gives the reason.

EXAMPLE 2.8. Let $A,B\in\mathbb{R}^{2\times 2}$, where

$$A = \begin{bmatrix} 2 & 2\\ 1 & 1\end{bmatrix},\quad B = \begin{bmatrix} 2 & x\\ 2 & 1\end{bmatrix},$$

and $x\in\mathbb{R}$. By some calculations, we obtain

$$\begin{aligned}&\|AB-BA\|_F^2 - \|A\odot B - B\odot A\|_F^2\\ &= (3x^2-20x+37)-(2x^2-8x+16) = x^2-12x+21.\end{aligned}$$

Thus, when $x=2$, we have 1, i.e., $\|AB-BA\|_F > \|A\odot B - B\odot A\|_F$; when $x=3$, we have -6, i.e., $\|AB-BA\|_F < \|A\odot B - B\odot A\|_F$.

3. Extension to tensor products

We first recall notations of tensors. An mth-order tensor

$$\mathscr{A} = [a_{i_1 i_2\ldots i_m}],\quad 1\leqslant i_s\leqslant n_s,\ 1\leqslant s\leqslant m,$$

is a multiway array consisting of $n_1 n_2\cdots n_m$ entries. We denote the set of all these tensors over the complex field $\mathbb{C}$ (or the real field $\mathbb{R}$) as $\mathbb{C}^{n_1\times n_2\times\cdots\times n_m}$ (or $\mathbb{R}^{n_1\times n_2\times\cdots\times n_m}$). If

$$n_1 = n_2 = \cdots = n_m = n,$$

then we call $\mathscr{A}$ an mth-order n-dimensional tensor. We denote the set of all complex (or real) mth-order n-dimensional tensors as $\mathbb{C}^{[m,n]}$ (or $\mathbb{R}^{[m,n]}$). The Frobenius norm of a tensor is defined by

$$\|\mathscr{A}\|_F := \left(\sum_{i_1,\dots,i_m=1}^{n_1,\dots,n_m} |a_{i_1\dots i_m}|^2 \right)^{\frac{1}{2}}, \quad \mathscr{A} = [a_{i_1\dots i_m}] \in \mathbb{C}^{n_1\times\cdots\times n_m},$$

where $|a_{i_1\dots i_m}|$ is the absolute value of $a_{i_1\dots i_m}$. The Kronecker product of Section 2.1 represents the classical product of two second-order tensors. In this section, we study the bound of the Frobenius norm of commutators based on two other tensor products: the contracted product and the T-product.

3.1. Contracted product

Let $\mathbf{c}$ and $\mathbf{r}$ be two subvectors of $\mathbf{m} = [1,2,\dots,m]$ and $\mathbf{d} = [1,2,\dots,d]$, respectively, and assume they have the same length. For tensors $\mathscr{A} \in \mathbb{C}^{[m,n]}$ and $\mathscr{B} \in \mathbb{C}^{[d,n]}$, the contracted product [3, 16, 17, 34] of $\mathscr{A}$ and $\mathscr{B}$ with respect to $\mathbf{c}$ and $\mathbf{r}$, is defined by

$$\mathscr{C} = \mathscr{A} \times_{\mathbf{c}}^{\mathbf{r}} \mathscr{B} \ \text{ with } \ \mathscr{C}(\mathbf{i}(\backslash\mathbf{c}),\mathbf{j}(\backslash\mathbf{r})) = \sum_{\{\mathbf{i}(\mathbf{c})|\mathbf{i}(\mathbf{c})=\mathbf{j}(\mathbf{r})\}} \mathscr{A}(\mathbf{i})\mathscr{B}(\mathbf{j}),$$

where $\mathbf{i} = [i_1,i_2,\dots,i_m]$, $\mathbf{j} = [j_1,j_2,\dots,j_d]$, $\mathbf{i}(\mathbf{c}) := [i_{c_1},i_{c_2},\dots,i_{c_s}]$ if $\mathbf{c} = [c_1,c_2,\dots,c_s]$, and $\backslash\mathbf{c}$ denotes the subvector whose elements arranged in ascending order are in $\mathbf{m}$ but not in $\mathbf{c}$.

In [37], Xie et al. concluded the following result of the commutator bound for some special contracted tensor products.

THEOREM 3.1. *Suppose $\mathscr{A},\mathscr{B} \in \mathbb{C}^{[m,n]}$. Let $\mathbf{c}$ and $\mathbf{r}$ be two subvectors of the vector $\mathbf{m} = [1,2,\dots,m]$ of the same length. If $\mathbf{r} = \rho(\mathbf{c})$ for some involution ρ (ρ is a permutation with ρ^2 being the identity) or if $\mathbf{r} = \sigma(\backslash\mathbf{c})$ for some permutation σ, then we have*

$$\|\mathscr{A} \times_{\mathbf{c}}^{\mathbf{r}} \mathscr{B} - \mathscr{B} \times_{\mathbf{c}}^{\mathbf{r}} \mathscr{A}\|_F \leqslant \sqrt{2}\|\mathscr{A}\|_F\|\mathscr{B}\|_F. \tag{3.1}$$

According to the proof in [37], we know that $(\mathscr{A},\mathscr{B})$ is a maximal pair for (3.1) if and only if one of the following conditions for the associated matricizations of the tensor contractions holds:

$\|\mathscr{A}_{\backslash\mathbf{c}\times\rho(\mathbf{c})}\mathscr{B}_{\mathbf{c}\times\backslash\mathbf{c}} - (\mathscr{B}_{\mathbf{c}\times\backslash\mathbf{c}})^T(\mathscr{A}_{\backslash\mathbf{c}\times\rho(\mathbf{c})})^T\|_F = \sqrt{2}\|\mathscr{A}_{\backslash\mathbf{c}\times\rho(\mathbf{c})}\|_F\|\mathscr{B}_{\mathbf{c}\times\backslash\mathbf{c}}\|_F$, when $\mathbf{r} = \rho(\mathbf{c})$;

$\|\mathscr{A}_{\sigma(\backslash\mathbf{c})\times\mathbf{c}}\mathscr{B}_{\sigma(\backslash\mathbf{c})\times\mathbf{c}} - \mathscr{B}_{\sigma(\backslash\mathbf{c})\times\mathbf{c}}\mathscr{A}_{\sigma(\backslash\mathbf{c})\times\mathbf{c}}\|_F = \sqrt{2}\|\mathscr{A}_{\sigma(\backslash\mathbf{c})\times\mathbf{c}}\|_F\|\mathscr{B}_{\sigma(\backslash\mathbf{c})\times\mathbf{c}}\|_F$, when $\mathbf{r} = \sigma(\backslash\mathbf{c})$.

Therefore, the problem of maximal pairs for (3.1) becomes the problem of maximal pairs for matrices; one can refer to [11, 15, 19, 28] for some sufficient and necessary conditions of maximal pairs for matrices.

EXAMPLE 3.2. We illustrate some examples that are maximal pairs for (3.1). For the case $\mathbf{r} = \rho(\mathbf{c})$, let $\mathbf{c} = [3,2]$, $\mathbf{r} = \rho(\mathbf{c}) = [2,3]$, and $\mathscr{A}, \mathscr{B} \in \mathbb{C}^{2\times 2\times 2}$, where all entries of the tensors $\mathscr{A}$ and $\mathscr{B}$ are zero except

$$\mathscr{A}(2,1,1) = -1, \quad \mathscr{B}(1,1,1) = 2.$$

Then we have

$$\begin{aligned}\|\mathscr{A} \times_{[3,2]}^{[2,3]} \mathscr{B} - \mathscr{B} \times_{[3,2]}^{[2,3]} \mathscr{A}\|_F &= \left\| \begin{bmatrix} 0 & 0 \\ -2 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix} \right\|_F \\ &= 2\sqrt{2} = \sqrt{2}\|\mathscr{A}\|_F\|\mathscr{B}\|_F.\end{aligned}$$

For the case $\mathbf{r} = \sigma(\backslash\mathbf{c})$, let $\mathbf{c} = [1,2]$, $\mathbf{r} = \sigma(\backslash\mathbf{c}) = [3,4]$, and suppose that all entries of $\mathscr{A}, \mathscr{B} \in \mathbb{C}^{2\times 2\times 2\times 2}$ are zero except

$$\mathscr{A}(2,1,1,1) = 1, \quad \mathscr{B}(1,1,2,1) = 1.$$

Therefore all entries of $\mathscr{C}_1 = \mathscr{A} \times_{[1,2]}^{[3,4]} \mathscr{B}$ and $\mathscr{C}_2 = \mathscr{B} \times_{[1,2]}^{[3,4]} \mathscr{A}$ are zero except

$$\mathscr{C}_1(1,1,1,1) = 1, \quad \mathscr{C}_2(2,1,2,1) = 1.$$

Again we have

$$\|\mathscr{A} \times_{[1,2]}^{[3,4]} \mathscr{B} - \mathscr{B} \times_{[1,2]}^{[3,4]} \mathscr{A}\|_F = \|\mathscr{C}_1 - \mathscr{C}_2\|_F = \sqrt{2} = \sqrt{2}\|\mathscr{A}\|_F\|\mathscr{B}\|_F.$$

We remark that if $\mathbf{c}$ and $\mathbf{r}$ are not linked as in Theorem 3.1, the norm of the commutator may exceed the value $\sqrt{2}$, as shown by the following example.

EXAMPLE 3.3. Let $\mathscr{A}, \mathscr{B} \in \mathbb{C}^{2\times 2\times 2}$, where

$$\mathscr{A}(:,:,1) = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \mathscr{A}(:,:,2) = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix},$$

$$\mathscr{B}(:,:,1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathscr{B}(:,:,2) = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

Suppose that $\mathbf{c} = [2,3]$ and $\mathbf{r} = [1,2]$. Then $\mathbf{r} \neq \rho(\mathbf{c})$ for any involution ρ and $\mathbf{r} \neq \sigma(\backslash\mathbf{c})$ for any permutation σ. We have

$$\begin{aligned}\|\mathscr{A} \times_{[2,3]}^{[1,2]} \mathscr{B} - \mathscr{B} \times_{[2,3]}^{[1,2]} \mathscr{A}\|_F &= \left\| \begin{bmatrix} 2 & -2 \\ 2 & -2 \end{bmatrix} - \begin{bmatrix} -1 & 2 \\ -1 & 1 \end{bmatrix} \right\|_F \\ &= \sqrt{43} > \sqrt{40} = \sqrt{2}\|\mathscr{A}\|_F\|\mathscr{B}\|_F.\end{aligned}$$

3.2. T-product

Next we introduce the definition of the T-product. For $\mathscr{A} \in \mathbb{R}^{m\times n\times p}$ and $\mathscr{B} \in \mathbb{R}^{n\times s\times p}$, we denote their frontal slices as $A^{(k)} = \mathscr{A}(:,:,k) \in \mathbb{R}^{m\times n}$ and $B^{(k)} = \mathscr{B}(:,:,k) \in \mathbb{R}^{n\times s}$, respectively, for $k = 1,2,\dots,p$. The operations bcirc, unfold, and fold are defined as follows [20, 23, 24]:

$$\text{bcirc}(\mathscr{A}) := \begin{bmatrix} A^{(1)} & A^{(p)} & A^{(p-1)} & \cdots & A^{(2)} \\ A^{(2)} & A^{(1)} & A^{(p)} & \cdots & A^{(3)} \\ \vdots & \vdots & & & \vdots \\ A^{(p)} & A^{(p-1)} & \cdots & A^{(2)} & A^{(1)} \end{bmatrix}, \quad \text{unfold}(\mathscr{A}) := \begin{bmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(p)} \end{bmatrix},$$

and $\text{fold}(\text{unfold}(\mathscr{A})) := \mathscr{A}$, which means that fold is the inverse opearator of unfold. We also denote the corresponding inverse operator as bcirc^{-1} such that

$$\text{bcirc}^{-1}(\text{bcirc}(\mathscr{A})) = \mathscr{A}.$$

DEFINITION 3.4. (T-product [20, 23, 24]) Let $\mathscr{A} \in \mathbb{R}^{m\times n\times p}$ and $\mathscr{B} \in \mathbb{R}^{n\times s\times p}$ be two real tensors. Then the T-product $\mathscr{A} * \mathscr{B}$ is an $m \times s \times p$ real tensor defined by

$$\mathscr{A} * \mathscr{B} := \text{fold}(\text{bcirc}(\mathscr{A})\text{unfold}(\mathscr{B})).$$

We remark that above concepts of the T-product could be easily extended to the complex field. Therefore we have the following norm inequality about two complex tensors.

THEOREM 3.5. *Let $\mathscr{A} \in \mathbb{C}^{n\times n\times p}$ and $\mathscr{B} \in \mathbb{C}^{n\times n\times p}$. Then we have*

$$\|\mathscr{A} * \mathscr{B} - \mathscr{B} * \mathscr{A}\|_F \leqslant \sqrt{2p}\|\mathscr{A}\|_F\|\mathscr{B}\|_F. \tag{3.2}$$

Proof. Based on the definition, we have

$$\|\text{bcirc}(\mathscr{A})\|_F = \left(p\|\mathscr{A}\|_F^2\right)^{\frac{1}{2}} = \sqrt{p}\|\mathscr{A}\|_F.$$

Observing the T-product, it is easy to see that $\text{unfold}(\mathscr{B})$ is the first column of $\text{bcirc}(\mathscr{B})$. Besides, the product of circulant matrices is circulant, which means that the block column matrix $\text{bcirc}(\mathscr{A})\text{unfold}(\mathscr{B})$ could determine the whole matrix $\text{bcirc}(\mathscr{A})\text{bcirc}(\mathscr{B})$ (see [33]). Thus,

$$\mathscr{A} * \mathscr{B} = \text{fold}(\text{bcirc}(\mathscr{A})\text{unfold}(\mathscr{B})) = \text{bcirc}^{-1}(\text{bcirc}(\mathscr{A})\text{bcirc}(\mathscr{B})).$$

Combining above two equalities and (1.1), we obtain

$$\begin{aligned} \|\mathscr{A} * \mathscr{B} - \mathscr{B} * \mathscr{A}\|_F &= \frac{1}{\sqrt{p}}\|\text{bcirc}(\mathscr{A} * \mathscr{B} - \mathscr{B} * \mathscr{A})\|_F \\ &= \frac{1}{\sqrt{p}}\|\text{bcirc}(\mathscr{A})\text{bcirc}(\mathscr{B}) - \text{bcirc}(\mathscr{B})\text{bcirc}(\mathscr{A})\|_F \\ &\leqslant \frac{1}{\sqrt{p}} \cdot \sqrt{2}\|\text{bcirc}(\mathscr{A})\|_F\|\text{bcirc}(\mathscr{B})\|_F = \sqrt{2p}\|\mathscr{A}\|_F\|\mathscr{B}\|_F. \qquad \square \end{aligned} \tag{3.3}$$

According to the proof of above theorem and the properties of the T-product, the characterization of maximal pairs for (3.2) can be derived.

COROLLARY 3.6. *Let* $n > 1$, $\mathscr{A} \in \mathbb{C}^{n\times n\times p}, \mathscr{B} \in \mathbb{C}^{n\times n\times p}$ *be two nonzero tensors, and* $A^{(q)} \in \mathbb{C}^{n\times n}$ *and* $B^{(q)} \in \mathbb{C}^{n\times n}$ *be their frontal slices for* $1 \leqslant q \leqslant p$. *Then* $(\mathscr{A}, \mathscr{B})$ *is a maximal pair for* (3.2) *if and only if*

(1) $\omega^{k(p-1)}A^{(1)} = \omega^{k(p-2)}A^{(2)} = \cdots = \omega^{k}A^{(p-1)} = A^{(p)}$,
$\omega^{k(p-1)}B^{(1)} = \omega^{k(p-2)}B^{(2)} = \cdots = \omega^{k}B^{(p-1)} = B^{(p)}$,
where $\omega = e^{-2\pi \mathrm{i}/p}$ *and* k *is an integer with* $0 \leqslant k \leqslant p-1$;

(2) $(A^{(1)}, B^{(1)})$ *is a maximal pair.*

Proof. Based on (3.3), we find that $(\mathscr{A}, \mathscr{B})$ is a maximal pair for (3.2) if and only if $(\mathrm{bcirc}(\mathscr{A}), \mathrm{bcirc}(\mathscr{B}))$ is maximal. As all circulant matrices can be diagonalized by the discrete Fourier matrix, so do block-circulant matrices. Then

$$(F_p \otimes I_n)\cdot \mathrm{bcirc}(\mathscr{A})\cdot (F_p^* \otimes I_n) = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_p \end{bmatrix} \equiv D_A$$

and

$$(F_p \otimes I_n)\cdot \mathrm{bcirc}(\mathscr{B})\cdot (F_p^* \otimes I_n) = \begin{bmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_p \end{bmatrix} \equiv D_B,$$

where F_p is the $p \times p$ discrete Fourier matrix and I_n is the $n \times n$ identity matrix [21, 24]. With this diagonalization, we have

$$\begin{aligned}
&\|\mathrm{bcirc}(\mathscr{A})\mathrm{bcirc}(\mathscr{B}) - \mathrm{bcirc}(\mathscr{B})\mathrm{bcirc}(\mathscr{A})\|_F^2 \\
&= \|D_A D_B - D_B D_A\|_F^2 \\
&= \|A_1B_1 - B_1A_1\|_F^2 + \cdots + \|A_pB_p - B_pA_p\|_F^2 \\
&\leqslant 2(\|A_1\|_F^2\|B_1\|_F^2 + \cdots + \|A_p\|_F^2\|B_p\|_F^2) \\
&\leqslant 2(\|A_1\|_F^2 + \cdots + \|A_p\|_F^2)(\|B_1\|_F^2 + \cdots + \|B_p\|_F^2) \\
&= 2\|D_A\|_F^2\|D_B\|_F^2 \\
&= 2\|\mathrm{bcirc}(\mathscr{A})\|_F^2\|\mathrm{bcirc}(\mathscr{B})\|_F^2.
\end{aligned}$$

Thus, $(\mathrm{bcirc}(\mathscr{A}), \mathrm{bcirc}(\mathscr{B}))$ is a maximal pair if and only if

(i) (A_i, B_i) is a maximal pair, for each $1 \leqslant i \leqslant p$;

(ii) $\|A_i\|_F^2\|B_j\|_F^2 = 0$, for any $i \neq j$, $1 \leqslant i, j \leqslant p$.

Since above properties are similar to those in the proof of Corollary 2.4, we find that only one diagonal block of D_A is nonzero, so is the corresponding block of D_B. Thus, $(\mathrm{bcirc}(\mathscr{A}),\mathrm{bcirc}(\mathscr{B}))$ is a maximal pair if and only if there exists an integer k with $0 \leqslant k \leqslant p-1$ such that

(a) (A_{k+1}, B_{k+1}) is maximal;

(b) $A_i = B_i = \mathbf{0}$, for any $i \neq k+1$, $1 \leqslant i \leqslant p$.

Next we want to show that the two properties (a) and (b) actually are equivalent to the sufficient and necessary conditions in the corollary. First, revealing the relationship between diagonal matrices of D_A and frontal slices of $\mathscr{A}$, we have

$$\begin{cases} A^{(1)} = \frac{1}{p}(\omega^0 A_1 + \omega^0 A_2 + \cdots + \omega^0 A_p) & = \frac{1}{p}\ \omega^0 A_{k+1}, \\ A^{(2)} = \frac{1}{p}(\omega^0 A_1 + \omega^1 A_2 + \cdots + \omega^{p-1} A_p) & = \frac{1}{p}\ \omega^k A_{k+1}, \\ \quad\cdots \\ A^{(p)} = \frac{1}{p}(\omega^0 A_1 + \omega^{p-1} A_2 + \cdots + \omega^{(p-1)(p-1)} A_p) & = \frac{1}{p}\ \omega^{k(p-1)} A_{k+1}. \end{cases}$$

Hence we have

$$\omega^{k(p-1)} A^{(1)} = \omega^{k(p-2)} A^{(2)} = \cdots = \omega^k A^{(p-1)} = A^{(p)}.$$

Similarly, $B^{(1)} = \frac{1}{p}\ \omega^0 B_{k+1}$ and

$$\omega^{k(p-1)} B^{(1)} = \omega^{k(p-2)} B^{(2)} = \cdots = \omega^k B^{(p-1)} = B^{(p)}.$$

Moreover, $(A^{(1)}, B^{(1)})$ is a maximal pair because (A_{k+1}, B_{k+1}) is a maximal pair.

Conversely, according to the relations between the frontal slices of $\mathscr{A}$, by some calculations, we obtain the following relationship between A_i and $A^{(i)}$ with $1 \leqslant i \leqslant p$:

$$A_i \ = \ \left[\omega^{(k-i+1)\cdot 0} + \omega^{(k-i+1)\cdot 1} + \cdots + \omega^{(k-i+1)(p-1)}\right] A^{(1)}.$$

Thus, $A_{k+1} = pA^{(1)}$ and $A_i = \mathbf{0}$ for other cases, which means that only A_{k+1} is nonzero. Similarly only $B_{k+1} = pB^{(1)}$ is nonzero. Furthermore, (A_{k+1}, B_{k+1}) is a maximal pair because $(A^{(1)}, B^{(1)})$ is a maximal pair. Therefore, the proof is completed. $\square$

EXAMPLE 3.7. We show an example that satisfies (3.2) with equality. Let $\mathscr{A}$ and $\mathscr{B}$ be $2 \times 2 \times 2$ real tensors, where

$$\mathrm{unfold}(\mathscr{A}) = \begin{bmatrix} A^{(1)} \\ A^{(2)} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} \ \text{and} \ \mathrm{unfold}(\mathscr{B}) = \begin{bmatrix} B^{(1)} \\ B^{(2)} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

It is easy to see that $A^{(1)} = A^{(2)}$, $B^{(1)} = B^{(2)}$, and $(A^{(1)}, B^{(1)})$ is a maximal pair. We have

$$\mathrm{unfold}(\mathscr{A} * \mathscr{B}) = \mathrm{bcirc}(\mathscr{A})\mathrm{unfold}(\mathscr{B}) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \\ 0 & 0 \\ 0 & 2 \end{bmatrix}$$

and

$$\operatorname{unfold}(\mathscr{B} * \mathscr{A}) = \operatorname{bcirc}(\mathscr{B})\operatorname{unfold}(\mathscr{A}) = \begin{bmatrix} 0&1&0&1 \\ 0&0&0&0 \\ 0&1&0&1 \\ 0&0&0&0 \end{bmatrix} \begin{bmatrix} 0&0 \\ 1&0 \\ 0&0 \\ 1&0 \end{bmatrix} = \begin{bmatrix} 2&0 \\ 0&0 \\ 2&0 \\ 0&0 \end{bmatrix}.$$

Hence

$$\|\mathscr{A} * \mathscr{B} - \mathscr{B} * \mathscr{A}\|_F = \sqrt{4 \cdot 2^2} = \sqrt{2 \cdot 2} \cdot \sqrt{2} \cdot \sqrt{2} = \sqrt{2p}\|\mathscr{A}\|_F \|\mathscr{B}\|_F.$$

It is interesting to notice that we may define an inner product for tensors in $\mathbb{R}^{n \times n \times p}$ just like

$$\langle \mathscr{A}, \mathscr{B} \rangle_{\overline{F}} := \operatorname{tr}\left(\operatorname{bcirc}(\mathscr{B})^* \operatorname{bcirc}(\mathscr{A})\right),$$

which means that we have another kind of norm for tensors:

$$\|\mathscr{A}\|_{\overline{F}}^2 = \langle \mathscr{A}, \mathscr{A} \rangle_{\overline{F}} = \operatorname{tr}\left(\operatorname{bcirc}(\mathscr{A})^* \operatorname{bcirc}(\mathscr{A})\right) = \|\operatorname{bcirc}(\mathscr{A})\|_F^2.$$

With this new norm, we have the following norm inequality.

COROLLARY 3.8. *Let $\mathscr{A} \in \mathbb{C}^{n \times n \times p}$ and $\mathscr{B} \in \mathbb{C}^{n \times n \times p}$. Then we have*

$$\|\mathscr{A} * \mathscr{B} - \mathscr{B} * \mathscr{A}\|_{\overline{F}} \leqslant \sqrt{2}\|\mathscr{A}\|_{\overline{F}} \|\mathscr{B}\|_{\overline{F}}. \tag{3.4}$$

Proof. Since $\operatorname{unfold}(\mathscr{A})$ is the first block column of the block-circulant matrix $\operatorname{bcirc}(\mathscr{A})$, we have

$$\|\mathscr{A}\|_{\overline{F}}^2 = \|\operatorname{bcirc}(\mathscr{A})\|_F^2 = p\|\operatorname{unfold}(\mathscr{A})\|_F^2 = p\|\mathscr{A}\|_F^2.$$

Thus, it directly follows from (3.2) that

$$\begin{aligned} \|\mathscr{A} * \mathscr{B} - \mathscr{B} * \mathscr{A}\|_{\overline{F}} &= \sqrt{p}\|\mathscr{A} * \mathscr{B} - \mathscr{B} * \mathscr{A}\|_F \\ &\leqslant \sqrt{2}p\|\mathscr{A}\|_F \|\mathscr{B}\|_F \\ &= \sqrt{2}\|\mathscr{A}\|_{\overline{F}} \|\mathscr{B}\|_{\overline{F}}. \quad \square \end{aligned}$$

In the proof, it is easy to see that this norm inequality is a scaled version of Theorem 3.5, which means that (3.2) and (3.4) are equivalent. Thus, the maximal pairs for (3.4) are the same as the ones for (3.2).

Acknowledgements. The authors would like to thank the editor and referees for their helpful suggestions and comments, which have considerably improved this paper.

REFERENCES

[1] P. ABSIL, R. MAHONY, R. SEPULCHRE, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, 2008.

[2] K. AUDENAERT, *Variance bounds, with an application to norm bounds for commutators*, Linear Algebra Appl., vol. 432 (2010), pp. 1126–1143.

[3] B. BADER, T. KOLDA, *Algorithm 862: MATLAB tensor classes for fast algorithm prototyping*, ACM Trans. Math. Software, vol. 32 (2006), pp. 635–653.

[4] R. BHATIA, *Matrix Analysis*, Springer, New York, 1997.

[5] R. BHATIA, F. KITTANEH, *Commutators, pinchings and spectral variation*, Oper. Matrices, vol. 2 (2008), pp. 143–151.

[6] A. BÖTTCHER, D. WENZEL, *How big can the commutator of two matrices be and how big is it typically?*, Linear Algebra Appl., vol. 403 (2005), pp. 216–228.

[7] A. BÖTTCHER, D. WENZEL, *Frobenius norm and the commutator*, Linear Algebra Appl., vol. 429 (2008), pp. 1864–1885.

[8] C. CHENG, K. FONG, W. LEI, *On some norm inequalities involving the commutator and $XY - YX^T$*, Linear Algebra Appl., vol. 438 (2013), pp. 2793–2807.

[9] C. CHENG, K. FONG, I. LOK, *Another proof for commutators with maximal Frobenius norm*, Recent Advances in Scientific Computing and Matrix Analysis, pp. 9–14, Editors: X. Jin, H. Sun, and S. Vong, International Press, 2011.

[10] C. CHENG, R. JIAO, *Proof of Wenzel's conjecture concerning singular values of the commutator of rank one matrices*, Linear Algebra Appl., vol. 592 (2020), pp. 165–174.

[11] C. CHENG, X. JIN, S. VONG, *A survey on the Böttcher-Wenzel conjecture and related problems*, Oper. Matrices, vol. 9 (2015), pp. 659–673.

[12] C. CHENG, C. LEI, *On Schatten p-norms of commutators*, Linear Algebra Appl., vol. 484 (2015), pp. 409–434.

[13] C. CHENG, Y. LIANG, *Some sharp bounds for the commutator of real matrices*, Linear Algebra Appl., vol. 521 (2017), pp. 263–282.

[14] C. CHENG, Y. LIANG, *Singular values, eigenvalues and diagonal elements of the commutator of 2×2 rank one matrices*, Electron. J. Linear Algebra, vol. 36 (2020), pp. 1–20.

[15] C. CHENG, S. VONG, D. WENZEL, *Commutators with maximal Frobenius norm*, Linear Algebra Appl., vol. 432 (2010), pp. 292–306.

[16] A. CICHOCKI, N. LEE, I. OSELEDETS, A. PHAN, Q. ZHAO, D. MANDIC, *Low-rank tensor networks for dimensionality reduction and large-scale optimization problems: perspectives and challenges part 1*, ArXiv preprint, arXiv:1609.00893v1, 2016.

[17] L. ELDÉN, B. SAVAS, *A Newton-Grassmann method for computing the best multilinear rank-(r_1,r_2,r_3) approximation of a tensor*, SIAM J. Matrix Anal. Appl., vol. 31 (2009), pp. 248–271.

[18] K. FONG, C. CHENG, I. LOK, *Another unitarily invariant norm attaining the minimum norm bound for commutators*, Linear Algebra Appl., vol. 433 (2010), pp. 1793–1797.

[19] K. FONG, I. LOK, C. CHENG, *A note on the norm of commutator and the norm of $XY - YX^T$*, Linear Algebra Appl., vol. 435 (2011), pp. 1191–1201.

[20] N. HAO, M. KILMER, K. BRAMAN, R. HOOVER, *Facial recognition using tensor-tensor decompositions*, SIAM J. Imaging Sci., vol. 6 (2013), pp. 437–463.

[21] X. JIN, S. VONG, *An Introduction to Applied Matrix Analysis*, Higher Education Press, Beijing, 2016.

[22] C. KHATRI, C. RAO, *Solutions to some functional equations and their applications to characterization of probability distributions*, Sankhya, vol. 30 (1968), pp. 167–180.

[23] M. KILMER, K. BRAMAN, N. HAO, R. HOOVER, *Third-order tensors as operators on matrices: a theoretical and computational framework with applications in imaging*, SIAM J. Matrix Anal. Appl., vol. 34 (2013), pp. 148–172.

[24] M. KILMER, C. MARTIN, *Factorization strategies for third-order tensors*, Linear Algebra Appl., vol. 435 (2011), pp. 641–658.

[25] F. KITTANEH, *Commutator inequalities associated with the polar decomposition*, Proc. Amer. Math. Soc., vol. 130 (2002), pp. 1279–1283.

[26] F. KITTANEH, *Inequalities for commutators of positive operators*, J. Funct. Anal., vol. 250 (2007), pp. 132–143.

[27] L. LÁSZLÓ, *Proof of Böttcher and Wenzel's conjecture on commutator norms for 3-by-3 matrices*, Linear Algebra Appl., vol. 422 (2007), pp. 659–663.
[28] I. LOK, *Norm inequalities for a matrix product analogous to the commutator*, M. Sc. Thesis, University of Macau, 2010.
[29] Z. LU, *Proof of the normal scalar curvature conjecture*, ArXiv preprint, arXiv:0711.3510v1 [math.DG], 2007.
[30] Z. LU, *Normal scalar curvature conjecture and its applications*, J. Funct. Anal., vol. 261 (2011), pp. 1284–1308.
[31] Z. LU, *Remarks on the Böttcher-Wenzel inequality*, Linear Algebra Appl., vol. 436 (2012), pp. 2531–2535.
[32] Z. LU, D. WENZEL, *Commutator estimates comprising the Frobenius norm – looking back and forth*, in: D. Bini, T. Ehrhardt, A. Karlovich, I. Spitkovsky (Eds.), *Large Truncated Toeplitz Matrices, Toeplitz Operators, and Related Topics*, in: Operator Theory: Advances and Applications, vol. 259, Birkhäuser, Cham, 2017, pp. 533–559.
[33] K. LUND, *The tensor t-function: a definition for functions of third-order tensors*, ArXiv preprint, arXiv:1806.07261, 2018.
[34] S. RAGNARSSON, C. VAN LOAN, *Block tensors unfoldings*, SIAM J. Matrix Anal. Appl., vol. 33 (2012), pp. 149–169.
[35] C. RAO, *Estimation of heteroscedastic variances in linear models*, Journal of the American Statistical Association, vol. 65 (1970), pp. 161–172.
[36] S. VONG, X. JIN, *Proof of Böttcher and Wenzel's conjecture*, Oper. Matrices, vol. 2 (2008), pp. 435–442.
[37] Z. XIE, C. CHENG, X. JIN, *Some norm inequalities for commutators of contracted tensor products*, Linear Algebra Appl., vol. 540 (2018), pp. 244–256.
[38] X. ZHAN, *Matrix Theory*, American Mathematical Society, Rhode Island, 2013.

(Received February 28, 2020)

Wei-Hui Liu
Department of Mathematics
University of Macau
Macao, China
e-mail: liu.weihui@connect.umac.mo

Ze-Jia Xie
Department of Mathematics
Shantou University
Shantou, 515063, China
e-mail: zjxie@stu.edu.cn

Xiao-qing Jin
Department of Mathematics
University of Macau
Macao, China
e-mail: xqjin@umac.mo