SIAM J. MATRIX ANAL. APPL.
Vol. 33, No. 1, pp. 149–169

BLOCK TENSOR UNFOLDINGS*

STEFAN RAGNARSSON† AND CHARLES F. VAN LOAN†

Abstract. Within the field of numerical multilinear algebra, block tensors are increasingly important. Accordingly, it is appropriate to develop an infrastructure that supports reasoning about block tensor computation. In this paper we establish concise notation that is suitable for the analysis and development of block tensor algorithms, prove several useful block tensor identities, and make precise the notion of a block tensor unfolding.

Key words. tensor, matrix computations, blocking

AMS subject classification. 15A69

DOI. 10.1137/110820609

1. Introduction. The field of matrix computations has matured to the point that it is not necessary to provide scalar-level verifications of basic block-level operations. For example, if

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^T \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

then without "ijk proof" it is *understood* that $C_{12} = A_{11}^T B_{12} + A_{21}^T B_{22}$ provided A and B are partitioned conformally. "Understandings" like this contribute to the culture of block matrix computations, enabling researchers to think at a high level when they are developing new algorithms and proofs.

It is our contention that the emerging field of tensor computations needs to develop a similar infrastructure that gracefully supports block tensor operations. By a *block tensor* we mean a tensor whose entries are themselves tensors. As with matrices, the act of blocking a tensor is the act of partitioning the index range vectors associated with each dimension. Thus, if $\mathcal{A} \in \mathbb{R}^{9\times5\times8}$ and

$$\begin{array}{rcl} 1{:}9 &=& \left[\begin{array}{cc|ccc|cccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array}\right] = \left[\begin{array}{c|c|c} 1{:}2 & 3{:}5 & 6{:}9 \end{array}\right], \\ 1{:}5 &=& \left[\begin{array}{ccc|cc} 1 & 2 & 3 & 4 & 5 \end{array}\right] = \left[\begin{array}{c|c} 1{:}3 & 4{:}5 \end{array}\right], \\ 1{:}8 &=& \left[\begin{array}{cc|cc|cc|cc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array}\right] = \left[\begin{array}{c|c|c|c} 1{:}2 & 3{:}4 & 5{:}6 & 7{:}8 \end{array}\right], \end{array} \tag{1.1}$$

then we are choosing to regard $\mathcal{A}$ as a 3-by-2-by-4 block tensor with block dimensions that are determined by the indicated partitionings of 1:9, 1:5, and 1:8. The colon notation can be used to specify the blocks. For example, the (2,1,3) block $\mathcal{A}_{213}$ is prescribed by $\mathcal{A}(3{:}5, 1{:}3, 5{:}6)$.

Block tensors are increasingly important for the same reasons that block matrices are increasingly important:

1. *Structure.* Block-level sparsity is a common pattern because of nearest-neighbor coupling and other reasons [15].

*Received by the editors January 10, 2011; accepted for publication (in revised form) by T. G. Kolda October 28, 2011; published electronically January 13, 2012. This work was supported in part by NSF contract DMS-1016284.
http://www.siam.org/journals/simax/33-1/82060.html
†Center for Applied Mathematics, Cornell University, Ithaca, NY 14853 (str23@cornell.edu, cv@cs.cornell.edu).

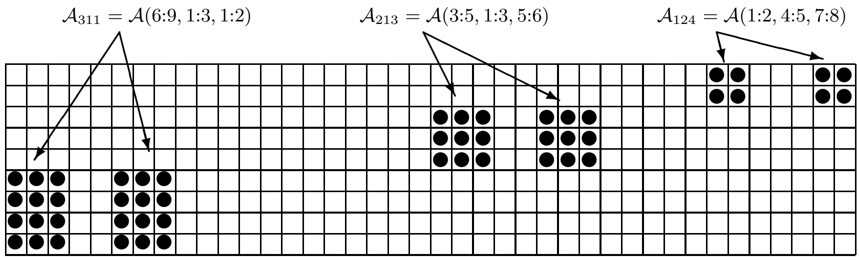


FIG. 1.1. *A vec-ordered, mode-*1 *unfolding of* $\mathcal{A} \in \mathbb{R}^{9\times 5\times 8}$ *with blocking* (1.1).

$$
\begin{array}{c}
2\,\{ \\ 3\,\{ \\ 4\,\{
\end{array}
\left[\begin{array}{c|c|c|c|c|c|c|c}
(\mathcal{A}_{111})_{(1)} & (\mathcal{A}_{121})_{(1)} & (\mathcal{A}_{112})_{(1)} & (\mathcal{A}_{122})_{(1)} & (\mathcal{A}_{113})_{(1)} & (\mathcal{A}_{123})_{(1)} & (\mathcal{A}_{114})_{(1)} & (\mathcal{A}_{124})_{(1)} \\ \hline
(\mathcal{A}_{211})_{(1)} & (\mathcal{A}_{221})_{(1)} & (\mathcal{A}_{212})_{(1)} & (\mathcal{A}_{222})_{(1)} & (\mathcal{A}_{213})_{(1)} & (\mathcal{A}_{223})_{(1)} & (\mathcal{A}_{214})_{(1)} & (\mathcal{A}_{224})_{(1)} \\ \hline
(\mathcal{A}_{311})_{(1)} & (\mathcal{A}_{321})_{(1)} & (\mathcal{A}_{312})_{(1)} & (\mathcal{A}_{322})_{(1)} & (\mathcal{A}_{313})_{(1)} & (\mathcal{A}_{323})_{(1)} & (\mathcal{A}_{314})_{(1)} & (\mathcal{A}_{324})_{(1)} \\ \hline
\underbrace{\qquad}_{6} & \underbrace{\quad}_{4} & \underbrace{\qquad}_{6} & \underbrace{\quad}_{4} & \underbrace{\qquad}_{6} & \underbrace{\quad}_{4} & \underbrace{\qquad}_{6} & \underbrace{\quad}_{4}
\end{array}\right]
$$

FIG. 1.2. *A "block vec"-ordered, mode-*1 *unfolding of* $\mathcal{A} \in \mathbb{R}^{9\times 5\times 8}$ *with blocking* (1.1).

2. *Generalization.* Block versions of point algorithms frequently have attractive features [14].
3. *Performance.* Blocking is the key to minimizing the overhead of communication [1].

Indeed, there is a very strong coupling between block tensor computations and block matrix computations. This is because the dominant paradigm for tensor computation involves the device of *unfolding.* An unfolded (or flattened) tensor is a matrix obtained by systematically reorganizing the tensor's entries into a two-dimensional array. In this framework, computations on a tensor $\mathcal{A}$ reduce to matrix computations on one or more of its unfoldings. For example, the higher-order singular value decomposition of a tensor involves computing the singular value decomposition of each modal unfolding [4]. See [13] for a nice overview of tensor decompositions and unfoldings.

Given all the advantages that result when a matrix computation is organized at the block level, it makes sense for an unfolding of a block tensor $\mathcal{A}$ to have a related block structure of its own. In particular, $\mathcal{A}$'s blocks should map to contiguous blocks in the unfolding. This is not the case when a typical "vec-oriented" unfolding is invoked [13]. Consider the mode-1 unfolding $\mathcal{A}_{(1)}$ of a 9-by-5-by-8 tensor $\mathcal{A}$ with blocking (1.1). The unfolding, which is displayed in Figure 1.1, is a 9-by-40 matrix whose ith row is $\mathrm{vec}(\mathcal{A}(i,:,:))^T$. (Recall that "vec-of-a-matrix" is the vector obtained by stacking its columns.)

Notice that in the unfolding, $\mathcal{A}$'s flattened blocks are not contiguous. The primary purpose of this paper is to show how to permute the rows and columns of a vec-oriented unfolding so that its blocks are unfoldings of the tensor blocks. An example of such an unfolding is displayed in Figure 1.2.

The paper is organized as follows. In section 2 we review well-known connections between vec(·), Kronecker products, transposition, and the perfect shuffle permutation. A block version of vec(·) is defined in section 3, and a related permutation is used to define the notion of a block unfolding. In section 4 we show how to formulate a tensor contraction as a block matrix multiplication using the tools developed.

2. Basic notation and operations. If $\mathcal{A} \in \mathbb{R}^{n_1\times\cdots\times n_d}$ and $\mathbf{i} = (i_1,\ldots,i_d)$, then $\mathcal{A}(\mathbf{i})$ denotes component $(i_1,\ldots,i_d)$ of tensor $\mathcal{A}$. We use calligraphic characters to designate tensors and bold lower case characters to denote vectors of integers. For $\mathcal{A}(\mathbf{i})$ to make sense we must have $1 \le i_k \le n_k$ for $k = 1{:}d$, i.e., $\mathbf{1} \le \mathbf{i} \le \mathbf{n}$. In general, if $\mathbf{i}$ and $\mathbf{j}$ have equal length, then $\mathbf{i} \le \mathbf{j}$ means that $i_k \le j_k$ for all k.

The MATLAB colon notation is used to specify index ranges. If $a < b$ and $c > 0$, then $a{:}b$ is the vector $[a, a+1, \ldots, b]$ and $a{:}c{:}b$ is the vector $[a, a+c, a+2c, \ldots, a+mc]$, where $m = \lfloor (b-a)/c \rfloor$, i.e., the largest integer that is less than or equal to $(b-a)/c$.

If $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times q}$, then the Kronecker product $A \otimes B \in \mathbb{R}^{mp\times nq}$ is the block matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}.$$

The outer product $\mathcal{C} = \mathcal{A} \circ \mathcal{B}$ of a tensor $\mathcal{A} \in \mathbb{R}^{j_1\times\cdots\times j_d}$ and a tensor $\mathcal{B} \in \mathbb{R}^{k_1\times\cdots\times k_e}$ is a tensor $\mathcal{C} \in \mathbb{R}^{j_1\times\cdots\times j_d\times k_1\times\cdots\times k_e}$ defined by

$$\mathcal{C}(\mathbf{i}) = \mathcal{A}(\mathbf{i}(1{:}d)) \cdot \mathcal{B}(\mathbf{i}(d+1{:}d+e)), \qquad \mathbf{1} \le \mathbf{i} \le [\,\mathbf{j}\ \mathbf{k}\,].$$

The order of $\mathcal{A} \circ \mathcal{B}$ is the order of $\mathcal{A}$ plus the order of $\mathcal{B}$. Note that $A \otimes B$ is an unfolding of the order-4 tensor $\mathcal{A} \circ \mathcal{B}$, where $\mathcal{A}$ and $\mathcal{B}$ are order-2 tensors (matrices) A and B.

2.1. The vec operation and ordering. If $\mathcal{A} \in \mathbb{R}^{n_1\times\cdots\times n_d}$ and $N = n_1 \cdots n_d$, then $\text{vec}(\mathcal{A}) \in \mathbb{R}^N$ is a column vector defined recursively by

$$\text{vec}(\mathcal{A}) \;=\; \begin{bmatrix} \text{vec}(\mathcal{A}^{(1)}) \\ \vdots \\ \text{vec}(\mathcal{A}^{(n_d)}) \end{bmatrix}, \tag{2.1}$$

where $\mathcal{A}^{(k)}$ is the order-$(d-1)$ tensor

$$\mathcal{A}^{(k)}(i_1,\ldots,i_{d-1}) \;=\; \mathcal{A}(i_1,\ldots,i_{d-1},k), \qquad 1 \le k \le n_d. \tag{2.2}$$

It is assumed that $\mathbf{1} \le \mathbf{i}(1{:}d-1) \le \mathbf{n}(1{:}d-1)$. If $d = 1$, then $\mathcal{A}$ is a column vector and $\text{vec}(\mathcal{A}) = \mathcal{A}$. If $d = 2$, then $\mathcal{A}$ is a matrix and $\text{vec}(\mathcal{A})$ stacks its columns. Each entry in tensor $\mathcal{A} \in \mathbb{R}^{n_1\times\cdots\times n_d}$ corresponds to a component of $\text{vec}(\mathcal{A})$. This implicitly defines an index mapping function $ivec(\,\cdot\,,\mathbf{n})$:

$$ivec(\mathbf{i},\mathbf{n}) \;=\; i_1 + (i_2-1)n_1 + (i_3-1)n_1n_2 + \cdots + (i_d-1)n_1\cdots n_{d-1}. \tag{2.3}$$

It is easy to show that if $v = \text{vec}(\mathcal{A})$, then

$$v_{ivec(\mathbf{i},\mathbf{n})} \;=\; \mathcal{A}(\mathbf{i}) \tag{2.4}$$

for all $\mathbf{i}$ that satisfy $\mathbf{1} \le \mathbf{i} \le \mathbf{n}$.

It should be noted that the "tensor vec" operation given by (2.1)–(2.4) reverts to the standard vec operation when $\mathcal{A}$ is a matrix [7].

2.2. Transposition, vec, Kronecker products, and permutation. There is an important connection between matrix transposition and perfect shuffle permutations [7, 8, 17, 18]. In particular, if $A \in \mathbb{R}^{q\times r}$ and $s = qr$, then

$$\mathrm{vec}(A^T) \;=\; \Pi_{q,r}^T \mathrm{vec}(A), \tag{2.5}$$

where $\Pi_{q,r} \in \mathbb{R}^{s\times s}$ is the (q,r) *perfect shuffle* permutation defined by

$$\Pi_{q,r} z \;=\; \begin{bmatrix} z(1{:}r{:}s) \\ z(2{:}r{:}s) \\ \vdots \\ z(r{:}r{:}s) \end{bmatrix}, \qquad z \in \mathbb{R}^s. \tag{2.6}$$

See [17]. If $Z \in \mathbb{R}^{r\times q}$ and $Y = Z^T$, then $\mathrm{vec}(Y) = \Pi_{q,r}\mathrm{vec}(Z)$. It is easy to verify that $\Pi_{q,r}^T = \Pi_{r,q}$.

If $f \in \mathbb{R}^q$ and $g \in \mathbb{R}^r$, then $g \otimes f$ is a perfect shuffle of $f \otimes g$:

$$\Pi_{q,r}\,(f \otimes g) \;=\; g \otimes f. \tag{2.7}$$

An important consequence of this result applies to the case when g is a block vector:

$$\mathrm{diag}(\Pi_{\rho_1,q},\ldots,\Pi_{\rho_\mu,q}) \cdot \Pi_{q,r} \cdot \left(f \otimes \begin{bmatrix} g_1 \\ \vdots \\ g_\mu \end{bmatrix} \right) \;=\; \begin{bmatrix} f \otimes g_1 \\ \vdots \\ f \otimes g_\mu \end{bmatrix}. \tag{2.8}$$

Here, $g_i \in \mathbb{R}^{\rho_i}$ and $r = \rho_1 + \cdots + \rho_\mu$.

Tensor transposition can also be characterized in terms of $\mathrm{vec}(\cdot)$ and perfect shuffles. If $\mathcal{A} \in \mathbb{R}^{n_1\times\cdots\times n_d}$ and $\mathbf{p}$ is a permutation of $1{:}d$, then $\mathcal{A}^{\langle \mathbf{p}\rangle} \in \mathbb{R}^{n_{p_1}\times\cdots\times n_{p_d}}$ denotes the $\mathbf{p}$-transpose of $\mathcal{A}$ and is defined by

$$\mathcal{A}^{\langle \mathbf{p}\rangle}(i_{p_1},\ldots,i_{p_d}) \;=\; \mathcal{A}(i_1,\ldots,i_d), \qquad \mathbf{1} \le \mathbf{i} \le \mathbf{n}, \tag{2.9}$$

i.e., $\mathcal{A}^{\langle \mathbf{p}\rangle}(\mathbf{i}(\mathbf{p})) = \mathcal{A}(\mathbf{i})$. The following lemma can be regarded as a generalization of (2.5).

LEMMA 2.1. *If* $\mathcal{A} \in \mathbb{R}^{N_1\times N_2\times N_3\times N_4}$ *and* $\mathcal{B} = \mathcal{A}^{\langle [\,1\;3\;2\;4\,]\rangle}$, *then*

$$\mathrm{vec}(\mathcal{B}) \;=\; (I_{N_4} \otimes \Pi_{N_3,N_2} \otimes I_{N_1})\mathrm{vec}(\mathcal{A}).$$

Proof. The proof follows from well-known facts that relate Kronecker products, $\mathrm{vec}(\cdot)$, and the perfect shuffle. See [7, 8, 17]. ∎

Although Lemma 2.1 addresses an order-4 transposition, the result can be applied to tensors of arbitrary order simply by "fusing" adjacent modes. For example, suppose $\mathcal{C} \in \mathbb{R}^{n_1\times\cdots\times n_7}$ and set $N_1 = n_1 n_2$, $N_2 = n_3$, $N_3 = n_4 n_5$, and $N_4 = n_6 n_7$. Define $\mathcal{A} \in \mathbb{R}^{N_1\times N_2\times N_3\times N_4}$ by

$$\mathcal{A}(j_1,j_2,j_3,j_4) \;=\; \mathcal{C}(\mathbf{i}), \quad \text{where} \quad \begin{cases} j_1 &=\; ivec(\mathbf{i}(1{:}2),\mathbf{n}(1{:}2)), \\ j_2 &=\; ivec(\mathbf{i}(3{:}3),\mathbf{n}(3{:}3)), \\ j_3 &=\; ivec(\mathbf{i}(4{:}5),\mathbf{n}(4{:}5)), \\ j_4 &=\; ivec(\mathbf{i}(6{:}7),\mathbf{n}(6{:}7)). \end{cases}$$

Observe that $\mathrm{vec}(\mathcal{A}) = \mathrm{vec}(\mathcal{C})$ and

$$(I_{N_4} \otimes \Pi_{N_3,N_2} \otimes I_{N_1})\mathrm{vec}(\mathcal{C}) = (I_{N_4} \otimes \Pi_{N_3,N_2} \otimes I_{N_1})\mathrm{vec}(\mathcal{A})$$
$$= \mathrm{vec}(\mathcal{A}^{\langle [\,1\,3\,2\,4\,] \rangle}) \;=\; \mathrm{vec}(\mathcal{C}^{\langle [\,1\,2\,4\,5\,3\,6\,7\,] \rangle}).$$

Two special applications of Lemma 2.1 are worth noting. Assume $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$. If $\mathbf{p} \;=\; [\,1{:}k{-}1\,,\; k{+}1\,,\; k\,,\; k{+}2{:}d\,]$, then

$$\mathrm{vec}(\mathcal{A}^{\langle \mathbf{p} \rangle}) \;=\; (I_{N_4} \otimes \Pi_{n_{k+1},n_k} \otimes I_{N_1})\mathrm{vec}(\mathcal{A}), \tag{2.10}$$

where $N_1 = n_1 \cdots n_{k-1}$ and $N_4 = n_{k+2} \cdots n_d$. This transposition swaps two adjacent modes, e.g.,

$$\mathcal{B} = \mathcal{A}^{\langle\, [1\,2\,4\,3\,5\,]\,\rangle} \Rightarrow \mathcal{A}(i_1, i_2, i_3, i_4, i_5) \;=\; \mathcal{B}(i_1, i_2, i_4, i_3, i_5).$$

On the other hand, if $\mathbf{p} \;=\; [\,k\,,\; 1{:}k-1\,,\; k+1{:}d\,]$, then

$$\mathrm{vec}(\mathcal{A}^{\langle \mathbf{p} \rangle}) \;=\; (I_{N_4} \otimes \Pi_{N_2,n_k})\mathrm{vec}(\mathcal{A}), \tag{2.11}$$

where $N_2 = n_1 \cdots n_{k-1}$ and $N_4 = n_{k+1} \cdots n_d$. This transposition "moves" a designated mode "to the front," e.g.,

$$\mathcal{B} = \mathcal{A}^{\langle\, [\,3\,1\,2\,4\,5\,]\,\rangle} \Rightarrow \mathcal{A}(i_1, i_2, i_3, i_4, i_5) \;=\; \mathcal{B}(i_3, i_1, i_2, i_4, i_5).$$

2.3. Unfolding a tensor. Converting a tensor to a matrix is an important operation in tensor computations [9, 10, 11, 13]. In order to unfold a tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ into a matrix, it is necessary to choose (a) an integer e that satisfies $1 \le e < d$ and (b) a permutation $\mathbf{p}$ of $1{:}d$. If

$$\mathbf{r} = \mathbf{p}(1{:}e), \tag{2.12}$$

$$\mathbf{c} = \mathbf{p}(e+1{:}d), \tag{2.13}$$

then the $\mathbf{r} \times \mathbf{c}$ *unfolding* of $\mathcal{A}$ is the matrix $\mathcal{A}_{\mathbf{r}\times\mathbf{c}}$ whose (α, β) entry is given by

$$\mathcal{A}_{\mathbf{r}\times\mathbf{c}}(\alpha, \beta) \;=\; \mathcal{A}^{\langle \mathbf{p} \rangle}(i_1, \ldots, i_e, j_1, \ldots, j_{d-e}), \tag{2.14}$$

where

$$\alpha = ivec(\mathbf{i}, \mathbf{n}(\mathbf{r})), \qquad \mathbf{1} \le \mathbf{i} \le \mathbf{n}(\mathbf{r}), \tag{2.15}$$

$$\beta = ivec(\mathbf{j}, \mathbf{n}(\mathbf{c})), \qquad \mathbf{1} \le \mathbf{j} \le \mathbf{n}(\mathbf{c}). \tag{2.16}$$

Note that $\mathcal{A}_{\mathbf{r}\times\mathbf{c}}$ has $n_{p_1} \cdots n_{p_e}$ rows and $n_{p_{e+1}} \cdots n_{p_d}$ columns. Each row and column of $\mathcal{A}_{\mathbf{r}\times\mathbf{c}}$ is the vec of a reduced-order subtensor. In particular, for all $\mathbf{i}$ and $\mathbf{j}$ that satisfy $\mathbf{1} \le \mathbf{i} \le \mathbf{n}(\mathbf{r})$ and $\mathbf{1} \le \mathbf{j} \le \mathbf{n}(\mathbf{c})$, we have

$$\mathcal{A}_{\mathbf{r}\times\mathbf{c}}(ivec(\mathbf{i}, \mathbf{n}(\mathbf{r})), :) = \mathrm{vec}(\,\mathcal{R}^{(\mathbf{i})})^T, \tag{2.17}$$

$$\mathcal{A}_{\mathbf{r}\times\mathbf{c}}(:, ivec(\mathbf{j}, \mathbf{n}(\mathbf{c}))) = \mathrm{vec}(\,\mathcal{C}^{(\mathbf{j})}), \tag{2.18}$$

where the tensors $\mathcal{R}^{(\mathbf{i})}$ and $\mathcal{C}^{(\mathbf{j})}$ are defined by

$$R^{(\mathbf{i})}(\mathbf{j}) = \mathcal{A}^{\langle \mathbf{p} \rangle}(i_1, \ldots, i_e, j_1, \ldots, j_{d-e}), \tag{2.19}$$

$$C^{(\mathbf{j})}(\mathbf{i}) = \mathcal{A}^{\langle \mathbf{p} \rangle}(i_1, \ldots, i_e, j_1, \ldots, j_{d-e}). \tag{2.20}$$

Especially important are the modal unfoldings. If $\mathbf{p} = [\,k\;\; 1{:}k{-}1\;\; k{+}1{:}d\,]$, then $\mathcal{A}_{\mathbf{r}\times\mathbf{c}}$ is a *mode-k unfolding* of $\mathcal{A}$. The columns of this matrix are referred to as *mode-e fibers* of $\mathcal{A}$. Special conventions are required if $\mathcal{A}$ is to be unfolded to either a column or a row vector. If $e = d$, then $\mathbf{c} = \emptyset$ and $\mathcal{A}_{\mathbf{r}\times\mathbf{c}} = \mathrm{vec}(\mathcal{A})$. Likewise, if $e = 0$, then $\mathbf{r} = \emptyset$ and $\mathcal{A}_{\mathbf{r}\times\mathbf{c}} = \mathrm{vec}(\mathcal{A})^T$.

2.4. Special cases. The preceding results take on a special form when $\mathcal{A}$ is a rank-1 tensor. Suppose $\mathcal{A} = a^{(1)} \circ \cdots \circ a^{(d)}$, where $a^{(k)} \in \mathbb{R}^{n_k}$ for $k = 1, \ldots, d$, i.e.,

$$\mathcal{A}(i_1, \ldots, i_d) = a^{(1)}(i_1) \cdots a^{(d)}(i_d), \qquad \mathbf{1} \le \mathbf{i} \le \mathbf{n}.$$

It follows from (2.1)–(2.4) that if

$$v = \text{vec}(a^{(1)} \circ \cdots \circ a^{(d)}),$$

then

$$v = a^{(d)} \otimes \cdots \otimes a^{(1)} \tag{2.21}$$

and

$$v_{ivec(\mathbf{i},\mathbf{n})} = a^{(1)}(i_1) \cdots a^{(d)}(i_d), \qquad \mathbf{1} \le \mathbf{i} \le \mathbf{n}. \tag{2.22}$$

If $\mathbf{p}$ is a permutation of $1{:}d$, then from the definition of the $\mathbf{p}$-transpose in (2.9) and the definition of $\mathcal{A}_{\mathbf{r}\times\mathbf{c}}$ in (2.12)–(2.16) we have

$$\mathcal{A}^{\langle \mathbf{p} \rangle} = a^{(p_1)} \circ \cdots \circ a^{(p_d)} \tag{2.23}$$

and

$$\mathcal{A}_{\mathbf{r}\times\mathbf{c}} = \text{vec}(a^{(r_1)} \circ \cdots \circ a^{(r_e)}) \cdot \text{vec}(a^{(c_1)} \circ \cdots \circ a^{(c_{d-e})})^T. \tag{2.24}$$

In other words, the unfolding of a rank-1 tensor is a rank-1 matrix. These rank-1 facts simplify some of the proofs that follow in the next section.

We consider another special case that relates to the multilinear product; see section 4.2. Suppose $\mathcal{B} = B^{(1)} \circ \cdots \circ B^{(d)}$, where $B^{(k)} \in \mathbb{R}^{q_k \times n_k}$ for $k = 1, \ldots, d$, i.e.,

$$\mathcal{B}(i_1, j_1, \ldots, i_d, j_d) = B^{(1)}(i_1, j_1) \cdots B^{(d)}(i_d, j_d).$$

Note that $\mathcal{B}$ is an order-$2d$ tensor. If $\mathbf{r} = 1{:}2{:}2d$, $\mathbf{c} = 2{:}2{:}2d$, and $\mathbf{p} = [\,\mathbf{r}\ \mathbf{c}\,]$, then for all $\mathbf{i}$ and $\mathbf{j}$ that satisfy $\mathbf{1} \le \mathbf{i} \le \mathbf{q}$ and $\mathbf{1} \le \mathbf{j} \le \mathbf{n}$ we have

$$\mathcal{B}_{\mathbf{r}\times\mathbf{c}}(\alpha, \beta) = B^{(1)}(i_1, j_1) \cdots B^{(d)}(i_d, j_d),$$

where $\alpha = ivec(\mathbf{i}, \mathbf{q})$ and $\beta = ivec(\mathbf{j}, \mathbf{n})$. However, this is precisely the (α, β) entry of the matrix $B^{(d)} \otimes \cdots \otimes B^{(1)}$. Thus,

$$\left(B^{(1)} \circ \cdots \circ B^{(d)}\right)_{[1:2:2d]\times[2:2:2d]} = B^{(d)} \otimes \cdots \otimes B^{(1)}. \tag{2.25}$$

3. Block notation and operations. In this section we formalize the notion of a block tensor [15], develop a block version of vec($\cdot$), and explain how to permute $\mathcal{A}_{\mathbf{r}\times\mathbf{c}}$ into a block matrix whose blocks are unfoldings of $\mathcal{A}$'s blocks. The presentation is simplified if we make use of multi-indexed subscripts. Suppose

$$\mathbf{1} \le \mathbf{i} \le \mathbf{s} = [s_1, \ldots, s_e], \qquad S = s_1 \cdots s_e,$$
$$\mathbf{1} \le \mathbf{j} \le \mathbf{t} = [t_1, \ldots, t_f], \qquad T = t_1 \cdots t_f.$$

To say that $v_{\mathbf{i}}$ is the $\mathbf{i}$th component of vector $v \in \mathbb{R}^S$ is to say that $v_{\mathbf{i}} = v_{ivec(\mathbf{i},\mathbf{s})}$. Similarly, if $D_1, \ldots, D_S$ are square matrices and $D = \text{diag}(\ldots, D_{\mathbf{i}}, \ldots)$, then D is a block diagonal matrix whose $\mathbf{i}$th diagonal block is $D_{ivec(\mathbf{i},\mathbf{s})}$. Finally, if $C = (C_{ij})$ is an S-by-T block matrix, then $C_{\mathbf{i},\mathbf{j}}$ is its $(\mathbf{i},\mathbf{j})$th block, i.e., $C_{\mathbf{i},\mathbf{j}} = C_{ivec(\mathbf{i},\mathbf{s}),ivec(\mathbf{j},\mathbf{t})}$.

3.1. Tensor blockings. We say that

$$\mathbf{M} \;=\; \{\mathbf{m}^{(1)}, \ldots, \mathbf{m}^{(d)}\} \tag{3.1}$$

is a *blocking* for $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ if

$$\mathbf{m}^{(k)} \;=\; \left[\, m_1^{(k)}, \ldots, m_{b_k}^{(k)} \,\right] \tag{3.2}$$

is a vector of positive integers that sums to n_k for $k = 1, \ldots, d$. If $\mathbf{1} \le \mathbf{i} \le \mathbf{b}$, then block $\mathbf{i}$ is the $m_{i_1}^{(1)} \times \cdots \times m_{i_d}^{(d)}$ tensor defined by

$$\mathcal{A}_{\mathbf{i}} \;=\; \mathcal{A}(\, \ell_{i_1}^{(1)}{:}u_{i_1}^{(1)}, \ldots, \ell_{i_d}^{(d)}{:}u_{i_d}^{(d)} \,), \tag{3.3}$$

where the lower and upper bound vectors $\boldsymbol{\ell}^{(1)}, \ldots, \boldsymbol{\ell}^{(d)}$ and $\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(d)}$ are defined by

$$\ell_j^{(k)} = m_1^{(k)} + \cdots + m_{j-1}^{(k)} \;+\; 1, \tag{3.4}$$

$$u_j^{(k)} = m_1^{(k)} + \cdots + m_{j-1}^{(k)} \;+\; m_j^{(k)} \tag{3.5}$$

for $k = 1, \ldots, d$. The blocking $\mathbf{M}$ identifies $\mathcal{A}$ as a $b_1 \times b_2 \times \cdots \times b_d$ block tensor. The number of elements in each tensor block $\mathcal{A}_{\mathbf{i}}$ turns out to be a quantity of importance and to that end we define the "volume function" $\mathrm{vol}_{\mathbf{M}}(\cdot)$ by

$$\mathrm{vol}_{\mathbf{M}}(\mathbf{i}) \;=\; m_{i_1}^{(1)} \cdots m_{i_d}^{(d)}, \qquad \mathbf{1} \le \mathbf{i} \le \mathbf{b}. \tag{3.6}$$

3.2. The $\mathbf{Vec_M(\cdot)}$ operation. If $\mathbf{M}$ is a blocking of $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ given by (3.1)–(3.5), then $\mathrm{vec}_{\mathbf{M}}(\mathcal{A})$ is the block vector

$$\mathrm{vec}_{\mathbf{M}}(\mathcal{A}) \;=\; \begin{bmatrix} v_{\mathbf{1}} \\ \vdots \\ v_{\mathbf{b}} \end{bmatrix}, \qquad v_{\mathbf{i}} \;=\; \mathrm{vec}(\mathcal{A}_{\mathbf{i}}), \tag{3.7}$$

where $\mathbf{1} \le \mathbf{i} \le \mathbf{b}$. In other words, $\mathrm{vec}_{\mathbf{M}}(\mathcal{A})$ stacks the vec's of $\mathcal{A}$'s blocks where the blocks are taken in the vec-order.

To illustrate this notation in the familiar matrix case, if

$$\mathbf{M} \;=\; \{m^{(1)}, m^{(2)}\} \;=\; \{[\, m_1^{(1)}\; m_2^{(1)} \,]\,,\, [\, m_1^{(2)}\; m_2^{(2)}\; m_3^{(2)} \,]\}$$

is a blocking for $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2}$, then we are choosing to regard $\mathcal{A}$ as a 2-by-3 block matrix

$$\mathcal{A} \;=\; \begin{array}{c} \left[\begin{array}{ccc} \mathcal{A}_{11} & \mathcal{A}_{12} & \mathcal{A}_{13} \\ \mathcal{A}_{21} & \mathcal{A}_{22} & \mathcal{A}_{23} \end{array}\right] \begin{array}{c} {\scriptstyle m_1^{(1)}} \\ {\scriptstyle m_2^{(1)}} \end{array} \\ \begin{array}{ccc} {\scriptstyle m_1^{(2)}} & {\scriptstyle m_2^{(2)}} & {\scriptstyle m_3^{(2)}} \end{array} \end{array} . \tag{3.8}$$

In this case, $\mathrm{vec}_{\mathbf{M}}(\cdot)$ and $\mathrm{vol}_{\mathbf{M}}(\cdot)$ are given by

$$\mathrm{vec}_{\mathbf{M}}(A) \;=\; \begin{bmatrix} v_{[1,1]} \\ v_{[2,1]} \\ v_{[1,2]} \\ v_{[2,2]} \\ v_{[1,3]} \\ v_{[2,3]} \end{bmatrix} = \begin{bmatrix} \mathrm{vec}(\mathcal{A}_{11}) \\ \mathrm{vec}(\mathcal{A}_{21}) \\ \mathrm{vec}(\mathcal{A}_{12}) \\ \mathrm{vec}(\mathcal{A}_{22}) \\ \mathrm{vec}(\mathcal{A}_{13}) \\ \mathrm{vec}(\mathcal{A}_{23}) \end{bmatrix}, \qquad \mathrm{vol}_{\mathbf{M}}(\mathbf{i}) \;=\; \begin{cases} m_1^{(1)} m_1^{(2)} & \text{if } \mathbf{i} = [1,1], \\ m_2^{(1)} m_1^{(2)} & \text{if } \mathbf{i} = [2,1], \\ m_1^{(1)} m_2^{(2)} & \text{if } \mathbf{i} = [1,2], \\ m_2^{(1)} m_2^{(2)} & \text{if } \mathbf{i} = [2,2], \\ m_1^{(1)} m_3^{(2)} & \text{if } \mathbf{i} = [1,3], \\ m_2^{(1)} m_3^{(2)} & \text{if } \mathbf{i} = [2,3]. \end{cases}$$

As we mentioned in the introduction, our goal is to permute the rows and columns of the unfolding $\mathcal{A}_{\mathbf{r}\times\mathbf{c}}$ so that its blocks are unfoldings of $\mathcal{A}$'s blocks. To be more precise, if $\mathcal{A} = (\mathcal{A}_{\mathbf{i}})$ is a block tensor, our goal is to determine permutation matrices $P_{\mathbf{R}}$ and $P_{\mathbf{C}}$ so that

$$\mathcal{A}_{\mathbf{R}\times\mathbf{C}} \;=\; P_{\mathbf{R}}\mathcal{A}_{\mathbf{r}\times\mathbf{c}}P_{\mathbf{C}}^T \tag{3.9}$$

is a block matrix whose blocks are the matrices $(\mathcal{A}_{\mathbf{k}})_{\mathbf{r}\times\mathbf{c}}$. It turns out that the permutations $P_{\mathbf{R}}$ and $P_{\mathbf{C}}$ map "vec-of-a-tensor" to "$\text{vec}_{\mathbf{M}}$-of-a-tensor." This is not surprising since the rows and columns of $\mathcal{A}_{\mathbf{r}\times\mathbf{c}}$ are vec's of reduced-order block tensors; see (2.17)–(2.20).

THEOREM 3.1. *Suppose* $\mathbf{M} = \{\mathbf{m}^{(1)},\ldots,\mathbf{m}^{(d)}\}$ *is a blocking of* $\mathcal{A}\in\mathbb{R}^{n_1\times\cdots\times n_d}$ *with*

$$\mathbf{m}^{(k)} = [m_1^{(1)},\ldots,m_{b_k}^{(k)}], \qquad k = 1,\ldots,d.$$

For $k = 1,\ldots,d$ *set*

$$N_k = n_1\cdots n_k,$$
$$\mathbf{M}_k = \{\mathbf{m}^{(1)},\ldots,\mathbf{m}^{(k)}\},$$

and define

$$Q_k \;=\; \begin{cases} I_{N_d} & \text{if } k = 1,\\ I_{N_d/N_k}\otimes\Gamma^{(k)} & \text{if } 1 < k \le d,\end{cases} \tag{3.10}$$

where $N_d/N_k = n_{k+1}n_{k+2}\cdots n_d$,

$$\Gamma^{(k)} \;=\; \text{diag}(\Gamma_1^{(k)},\ldots,\Gamma_{b_k}^{(k)}), \tag{3.11}$$

and

$$\Gamma_j^{(k)} \;=\; \text{diag}(\ldots,\Pi_{\text{vol}_{\mathbf{M}_{k-1}}(\mathbf{i}),m_j^{(k)}},\ldots)\cdot\Pi_{m_j^{(k)},N_{k-1}}, \qquad \mathbf{1}\le\mathbf{i}\le\mathbf{b}(1{:}k-1). \tag{3.12}$$

The permutation matrix $P_{\mathbf{M}}$ *defined by*

$$P_{\mathbf{M}} \;=\; Q_d\cdots Q_2Q_1$$

has the property that

$$\text{vec}_{\mathbf{M}}(\mathcal{A}) \;=\; P_{\mathbf{M}}\text{vec}(\mathcal{A}).$$

Proof. Since both $\text{vec}(\cdot)$ and $\text{vec}_{\mathbf{M}}(\cdot)$ are linear operators and any tensor is the sum of rank-1 tensors, it suffices to prove the theorem for the case

$$\mathcal{A} \;=\; a^{(1)}\circ\cdots\circ a^{(d)},$$

where each $a^{(k)}\in\mathbb{R}^{n_k}$ is blocked as follows:

$$a^{(k)} \;=\; \begin{bmatrix} a_1^{(k)}\\ \vdots\\ a_{b_k}^{(k)}\end{bmatrix}\begin{matrix}\}\,m_1^{(k)},\\ \\ \}\,m_{b_k}^{(k)}.\end{matrix}$$

We proceed by induction noting that the theorem is true if $d = 1$ because in that case, $\text{vec}_{\mathbf{M}}(\mathcal{A}) = \text{vec}(\mathcal{A})$. Assume that the theorem holds for block tensors with order $d-1$ or less with $d > 1$. Define

$$\begin{aligned}\widehat{\mathcal{A}} &= a^{(1)} \circ \cdots \circ a^{(d-1)},\\ \widehat{\mathbf{M}} &= \mathbf{M}_{d-1},\\ \widehat{\mathbf{b}} &= \mathbf{b}(1{:}d-1),\end{aligned}$$

and observe that $\widehat{\mathbf{M}}$ is a blocking for $\widehat{\mathcal{A}}$, an order-$(d-1)$ tensor. It follows by induction that

$$\text{vec}_{\widehat{\mathbf{M}}}(\widehat{\mathcal{A}}\,) \;=\; P_{\widehat{\mathbf{M}}}\,\text{vec}(\widehat{\mathcal{A}}\,). \tag{3.13}$$

From the definition of $\text{vec}_{\mathbf{M}}(\cdot)$ in (3.7), we have

$$\text{vec}_{\widehat{\mathbf{M}}}(\widehat{\mathcal{A}}\,) \;=\; \begin{bmatrix} v_{\mathbf{1}} \\ \vdots \\ v_{\widehat{\mathbf{b}}} \end{bmatrix}, \qquad v_{\mathbf{i}} = a^{(d-1)}_{i_{d-1}} \otimes \cdots \otimes a^{(1)}_{i_1} \tag{3.14}$$

for all $\mathbf{i}$ that satisfy $\mathbf{1} \le \mathbf{i} \le \widehat{\mathbf{b}}$. Equation (2.21) says that

$$\text{vec}(\mathcal{A}) \;=\; a^{(d)} \otimes (a^{(d-1)} \otimes \cdots \otimes a^{(1)}) \;=\; a^{(d)} \otimes \text{vec}(\widehat{\mathcal{A}}\,),$$

and so

$$(I_{n_d} \otimes P_{\widehat{\mathbf{M}}})\text{vec}(\mathcal{A}) \;=\; a^{(d)} \otimes v \;=\; \begin{bmatrix} a^{(d)}_1 \\ \vdots \\ a^{(d)}_{b_d} \end{bmatrix} \otimes v \;=\; \begin{bmatrix} a^{(d)}_1 \otimes v \\ \vdots \\ a^{(d)}_{b_d} \otimes v \end{bmatrix}. \tag{3.15}$$

Using (2.8), we have for $j = 1, \ldots, b_d$ that

$$\Gamma^{(d)}_j \left(a^{(d)}_j \otimes v \right) \;=\; \begin{bmatrix} a^{(d)}_j \otimes v_{\mathbf{1}} \\ \vdots \\ a^{(d)}_j \otimes v_{\widehat{\mathbf{b}}} \end{bmatrix},$$

where

$$\Gamma^{(d)}_j \;=\; \text{diag}\left(\Pi_{\text{vol}_{\widehat{\mathbf{M}}}(\mathbf{1}), m^{(d)}_j}, \ldots, \Pi_{\text{vol}_{\widehat{\mathbf{M}}}(\mathbf{b}(1:d-1)), m^{(d)}_j}\right) \cdot \Pi_{m^{(d)}_j, N/n_d}.$$

Thus, if $\Gamma^{(d)} = \text{diag}(\Gamma^{(d)}_1, \ldots, \Gamma^{(d)}_{b_d})$, then

$$\Gamma^{(d)} \left[\begin{array}{c} a^{(d)}_1 \otimes v \\ \hline \vdots \\ \hline a^{(d)}_{b_d} \otimes v \end{array}\right] \;=\; \left[\begin{array}{c} a^{(d)}_1 \otimes v_{\mathbf{1}} \\ \vdots \\ a^{(d)}_1 \otimes v_{\widehat{\mathbf{b}}} \\ \hline \vdots \\ \hline a^{(d)}_{b_d} \otimes v_{\mathbf{1}} \\ \vdots \\ a^{(d)}_{b_d} \otimes v_{\widehat{\mathbf{b}}} \end{array}\right] \;=\; \text{vec}_{\mathbf{M}}(\mathcal{A}). \tag{3.16}$$

Combining this equation with (3.15), we have

$$\Gamma^{(d)}(I_{n_d} \otimes P_{\widehat{\mathbf{M}}})\mathrm{vec}(\mathcal{A}) \;=\; \mathrm{vec}_{\mathbf{M}}(\mathcal{A}),$$

and so $P_{M} = \Gamma^{(d)}(I_{n_d} \otimes P_{\widehat{\mathbf{M}}})$. But by induction

$$P_{\widehat{\mathbf{M}}} \;=\; \widehat{Q}_{d-1} \cdots \widehat{Q}_2\, \widehat{Q}_1,$$

where

$$\widehat{Q}_k \;=\; \begin{cases} I_{N_{d-1}} & \text{if } k = 1, \\ I_{N_{d-1}/N_k} \otimes \Gamma^{(k)} & \text{if } 1 < k \le d-1. \end{cases}$$

It follows that

$$\begin{aligned} P_{\mathbf{M}} &= \Gamma^{(d)}(I_{n_d} \otimes P_{\widehat{\mathbf{M}}}) \\ &= \Gamma^{(d)}(I_{n_d} \otimes \widehat{Q}_{d-1}) \cdots (I_{n_d} \otimes \widehat{Q}_2)(I_{n_d} \otimes \widehat{Q}_1) \\ &= (I_{N_d/N_d} \otimes \Gamma^{(d)})(I_{N_d/N_{d-1}} \otimes \Gamma^{(d-1)}) \cdots (I_{N_d/N_2} \otimes \Gamma^{(2)})(I_{N_d}) \\ &= Q_d\, Q_{d-1} \cdots Q_2\, Q_1, \end{aligned}$$

completing the proof. ∎

The permutation $P_{\mathbf{M}}$ has a particularly simple form if the blocking is uniform in each dimension.

Corollary 3.2. *Suppose* $\mathbf{M}$ *is defined by* (3.1)–(3.5)*. If*

$$\begin{aligned} m_1^{(k)} = \cdots = m_{b_k}^{(k)} &\;=\; \mu_k, \\ N_k &= n_1 \cdots n_k, \\ B_k &= b_1 \cdots b_k, \\ D_k &= \mu_1 \cdots \mu_k \end{aligned}$$

for $k = 1, \ldots, d$*, then* $P_{M} = Q_d \cdots Q_2 Q_1$*, where*

$$Q_k \;=\; \begin{cases} I_{N_d} & \text{if } k = 1, \\ I_{b_k N_d/N_k} \otimes \Pi_{\mu_k, B_{k-1}} \otimes I_{D_{k-1}} & \text{if } 1 < k \le d. \end{cases}$$

Proof. Observe that $\mathrm{vol}_{\mathbf{M}_{k-1}}(\mathbf{i}) \;=\; \mu_1 \cdots \mu_{k-1}$. It follows from the definition of $\Gamma_j^{(k)}$ in (3.12) that

$$\Gamma_j^{(k)} \;=\; \left(I_{B_{k-1}} \otimes \Pi_{D_{k-1}, \mu_k}\right) \Pi_{\mu_k, N_{k-1}}.$$

Using the well-known Kronecker product identity

$$(I_s \otimes \Pi_{r,q})\, \Pi_{q,rs} \;=\; \Pi_{q,s} \otimes I_r,$$

it follows that

$$\Gamma_j^{(k)} \;=\; \Pi_{\mu_k, B_{k-1}} \otimes I_{D_{k-1}}.$$

See [17]. From (3.10) we have

$$\Gamma^{(k)} \;=\; I_{b_k} \otimes \Pi_{\mu_k, B_{k-1}} \otimes I_{D_{k-1}},$$

and so

$$Q_k \;=\; I_{N_d/N_k} \otimes \Gamma^{(k)} \;=\; I_{N_d b_k/N_k} \otimes \Pi_{\mu_k, B_{k-1}} \otimes I_{D_{k-1}}.$$

This completes the proof. ☐

It is interesting to note that the transition from $\mathrm{vec}(\mathcal{A})$ to $\mathrm{vec}_{\mathbf{M}}(\mathcal{A})$ via the sequence

$$Q_2 \cdot \mathrm{vec}(\mathcal{A}) \;\to\; Q_3 \cdot (Q_2 \cdot \mathrm{vec}(\mathcal{A})) \;\to \cdots \to\; Q_d \cdot (Q_{d-1} \cdots Q_2 \cdot \mathrm{vec}(\mathcal{A}))$$

is actually a sequence of transpositions. To illustrate this, we assume $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3 \times n_4}$ and define the order-8 tensor $\mathcal{A}^{(1)}$ by

$$\mathcal{A}(i_1, i_2, i_3, i_4) \;=\; \mathcal{A}^{(1)}(\delta_1, \beta_1, \delta_2, \beta_2, \delta_3, \beta_3, \delta_4, \beta_4),$$

where $\mathbf{1} \le \mathbf{i} \le \mathbf{n}$ and the δ_k and β_k are uniquely defined by

$$i_k = \delta_k + (\beta_k - 1) b_k, \qquad 1 \le \delta_k \le \mu_k.$$

This says that $\mathcal{A}^{(1)} \in \mathbb{R}^{\mu_1 \times b_1 \times \mu_2 \times b_2 \times \mu_3 \times b_3 \times \mu_4 \times b_4}$. In the $d = 4$ case, the Q-matrices in Corollary 4.2 are given by

$$\begin{aligned} Q_2 &= I_{b_2 n_3 n_4} \otimes \Pi_{\mu_2, b_1} \otimes I_{\mu_1}, \\ Q_3 &= I_{b_3 n_4} \otimes \Pi_{\mu_3, b_1 b_2} \otimes I_{\mu_1 \mu_2}, \\ Q_4 &= I_{b_4} \otimes \Pi_{\mu_4, b_1 b_2 b_3} \otimes I_{\mu_1 \mu_2 \mu_3}. \end{aligned}$$

Note from Lemma 2.1 that these permutations correspond to transpositions. Indeed, if we define the tensors $\mathcal{A}^{(2)}$, $\mathcal{A}^{(3)}$, $\mathcal{A}^{(4)}$ by

$$\left.\begin{array}{l} \mathcal{A}^{(2)}(\delta_1, \delta_2, \beta_1, \beta_2, \delta_3, \beta_3, \delta_4, \beta_4) \\ \mathcal{A}^{(3)}(\delta_1, \delta_2, \delta_3, \beta_1, \beta_2, \beta_3, \delta_4, \beta_4) \\ \mathcal{A}^{(4)}(\delta_1, \delta_2, \delta_3, \delta_4, \beta_1, \beta_2, \beta_3, \beta_4) \end{array}\right\} \;=\; \mathcal{A}^{(1)}(\delta_1, \beta_1, \delta_2, \beta_2, \delta_3, \beta_3, \delta_4, \beta_4),$$

then it can be shown via Lemma 2.1 that

$$\begin{aligned} \mathrm{vec}(\mathcal{A}^{(1)}) &= Q_1 \mathrm{vec}(\mathcal{A}) \;=\; \mathrm{vec}(\mathcal{A}), \\ \mathrm{vec}(\mathcal{A}^{(2)}) &= Q_2 \mathrm{vec}(\mathcal{A}^{(1)}), \\ \mathrm{vec}(\mathcal{A}^{(3)}) &= Q_3 \mathrm{vec}(\mathcal{A}^{(2)}), \\ \mathrm{vec}_{\mathbf{M}}(\mathcal{A}) \;=\; \mathrm{vec}(\mathcal{A}^{(4)}) &= Q_4 \mathrm{vec}(\mathcal{A}^{(3)}). \end{aligned}$$

Thus, the order-8 tensor $\mathcal{A}^{(4)}$ has the property that $\mathrm{vec}(\mathcal{A}^{(4)}) = \mathrm{vec}_{\mathbf{M}}(\mathcal{A})$. Moreover, $\mathcal{A}(\mathbf{i}) = \mathcal{A}_{\boldsymbol{\beta}}(\boldsymbol{\delta})$, showing that entry $\mathbf{i}$ is entry $\boldsymbol{\delta}$ of block $\boldsymbol{\beta}$.

3.3. Block unfoldings. We now specify the permutation matrices $P_{\mathbf{R}}$ and $P_{\mathbf{C}}$ in (3.9) that turn $\mathcal{A}_{\mathbf{r}\times\mathbf{c}}$ into a block matrix with block entries that are $\mathbf{r}\times\mathbf{c}$ unfoldings of $\mathcal{A}$'s blocks.

THEOREM 3.3. *Suppose* $\mathbf{M} = \{\mathbf{m}^{(1)}, \ldots, \mathbf{m}^{(d)}\}$ *is a blocking of* $\mathcal{A} \in \mathbb{R}^{n_1\times\cdots\times n_d}$ *with*

$$\mathbf{m}^{(k)} \;=\; [m_1^{(k)}, \ldots, m_{b_k}^{(k)}], \qquad k = 1, \ldots, d.$$

Let e *be an integer that satisfies* $1 \le e < d$, *and assume that* $\mathbf{p}$ *is a permutation of* $1{:}d$. *Define*

$$\begin{array}{lllllllll} \mathbf{r} & = & \mathbf{p}(1{:}e), & \mathbf{R} & = & \{\mathbf{m}^{(r_1)}, \ldots, \mathbf{m}^{(r_e)}\}, & B_{rows} & = & b_{r_1}\cdots b_{r_e}, \\ \mathbf{c} & = & \mathbf{p}(e+1{:}d), & \mathbf{C} & = & \{\mathbf{m}^{(c_1)}, \ldots, \mathbf{m}^{(c_{d-e})}\}, & B_{cols} & = & b_{c_1}\cdots b_{c_{d-e}}. \end{array}$$

The matrix

$$\mathcal{A}_{\mathbf{R}\times\mathbf{C}} \;=\; P_{\mathbf{R}}\mathcal{A}_{\mathbf{r}\times\mathbf{c}}P_{\mathbf{C}}^T$$

is a B_{rows}*-by-*B_{cols} *block matrix whose block entries are specified by*

$$(\mathcal{A}_{\mathbf{R}\times\mathbf{C}})_{\mathbf{k}(\mathbf{r}),\mathbf{k}(\mathbf{c})} \;=\; (\mathcal{A}_{\mathbf{k}})_{\mathbf{r}\times\mathbf{c}}\,, \qquad \mathbf{1} \le \mathbf{k} \le \mathbf{b}. \tag{3.17}$$

That is to say, if $\mu = ivec(\mathbf{k}(\mathbf{r}), \mathbf{b}(\mathbf{r}))$ *and* $\tau = ivec(\mathbf{k}(\mathbf{c}), \mathbf{b}(\mathbf{c}))$, *then the* (μ, τ) *block of* $\mathcal{A}_{\mathbf{R}\times\mathbf{C}}$ *is the* $\mathbf{r}\times\mathbf{c}$ *unfolding of the* $\mathbf{k}$*th block of* $\mathcal{A}$.

Proof. By linearity there is no loss of generality in assuming that

$$\mathcal{A} \;=\; a^{(1)} \circ \cdots \circ a^{(d)},$$

where each $a^{(k)} \in \mathbb{R}^{n_k}$ is blocked as follows:

$$a^{(k)} \;=\; \left[\begin{array}{c} a_1^{(k)} \\ \vdots \\ a_{b_k}^{(k)} \end{array}\right] \begin{array}{l} \}\, m_1^{(k)} \\ \\ \}\, m_{b_k}^{(k)} \end{array}.$$

From (2.24) we know that

$$\mathcal{A}_{\mathbf{r}\times\mathbf{c}} \;=\; \mathrm{vec}(\, a^{(r_1)} \circ \cdots \circ a^{(r_e)}\,) \cdot \mathrm{vec}(\, a^{(c_1)} \circ \cdots \circ a^{(c_{d-e})}\,)^T.$$

Since $\mathbf{R}$ is a blocking for $a^{(r_1)} \circ \cdots \circ a^{(r_e)}$ and $\mathbf{C}$ is a blocking for $a^{(c_1)} \circ \cdots \circ a^{(c_{d-e})}$, it follows from Theorem 3.1 that

$$P_{\mathbf{R}}A_{\mathbf{r}\times\mathbf{c}}P_{\mathbf{C}}^T \;=\; yz^T,$$

where $y \;=\; \mathrm{vec}_{\mathbf{R}}(a^{(r_1)} \circ \cdots \circ a^{(r_e)})$ and $z \;=\; \mathrm{vec}_{\mathbf{C}}(a^{(c_1)} \circ \cdots \circ a^{(c_{d-e})})$. These block vectors are specified by

$$y = \left[\begin{array}{c} y_{\mathbf{1}} \\ \vdots \\ y_{\mathbf{b}(\mathbf{r})} \end{array}\right], \qquad y_{\mathbf{i}} \;=\; \mathrm{vec}(a_{i_1}^{(r_1)} \circ \cdots \circ a_{i_e}^{(r_e)}), \qquad \mathbf{1} \le \mathbf{i} \le \mathbf{b}(\mathbf{r}), \tag{3.18}$$

$$z = \left[\begin{array}{c} z_{\mathbf{1}} \\ \vdots \\ z_{\mathbf{b}(\mathbf{c})} \end{array}\right], \qquad z_{\mathbf{j}} \;=\; \mathrm{vec}(a_{j_1}^{(c_1)} \circ \cdots \circ a_{j_{d-e}}^{(c_{d-e})}), \qquad \mathbf{1} \le \mathbf{j} \le \mathbf{b}(\mathbf{c}), \tag{3.19}$$

and so the $(\mathbf{i},\mathbf{j})$th block of $\mathcal{A}_{\mathbf{R}\times\mathbf{C}}$ is given by

$$(\mathcal{A}_{\mathbf{R}\times\mathbf{C}})_{\mathbf{i},\mathbf{j}} \;=\; y_{\mathbf{i}}\, z_{\mathbf{j}}^T. \tag{3.20}$$

On the other hand, from (3.17)

$$\begin{aligned}(\mathcal{A}_{\mathbf{k}})_{\mathbf{r}\times\mathbf{c}} &= \left(a_{k_1}^{(1)} \circ \cdots \circ a_{k_d}^{(d)}\right)_{\mathbf{r}\times\mathbf{c}} \\ &= \operatorname{vec}\left(a_{k_{r_1}}^{(r_1)} \circ \cdots \circ a_{k_{r_e}}^{(r_e)}\right) \cdot \operatorname{vec}\left(a_{k_{c_1}}^{(c_1)} \circ \cdots \circ a_{k_{c_{d-e}}}^{(c_{d-e})}\right)^T.\end{aligned}$$

It follows from (3.18)–(3.20) that if $\mathbf{i} = \mathbf{k}(\mathbf{r})$ and $\mathbf{j} = \mathbf{k}(\mathbf{c})$, then

$$(\mathcal{A}_{\mathbf{k}})_{\mathbf{r}\times\mathbf{c}} \;=\; y_{\mathbf{i}}\, z_{\mathbf{j}}^T \;=\; (\mathcal{A}_{\mathbf{R}\times\mathbf{C}})_{\mathbf{i},\mathbf{j}},$$

which completes the proof. ∎

To illustrate the theorem, suppose $\mathcal{A}$ is a 2-by-4-by-3-by-2 block tensor. If $\mathbf{r} = [1\ 3]$ and $\mathbf{c} = [2\ 4]$, then

$$\mathcal{A}_{\mathbf{R}\times\mathbf{C}} \;=\; \begin{array}{cc} \left[\begin{array}{cccccccc} \widetilde{\mathcal{A}}_{1111} & \widetilde{\mathcal{A}}_{1211} & \widetilde{\mathcal{A}}_{1311} & \widetilde{\mathcal{A}}_{1411} & \widetilde{\mathcal{A}}_{1112} & \widetilde{\mathcal{A}}_{1212} & \widetilde{\mathcal{A}}_{1312} & \widetilde{\mathcal{A}}_{1412} \\ \widetilde{\mathcal{A}}_{2111} & \widetilde{\mathcal{A}}_{2211} & \widetilde{\mathcal{A}}_{2311} & \widetilde{\mathcal{A}}_{2411} & \widetilde{\mathcal{A}}_{2112} & \widetilde{\mathcal{A}}_{2212} & \widetilde{\mathcal{A}}_{2312} & \widetilde{\mathcal{A}}_{2412} \\ \widetilde{\mathcal{A}}_{1121} & \widetilde{\mathcal{A}}_{1221} & \widetilde{\mathcal{A}}_{1321} & \widetilde{\mathcal{A}}_{1421} & \widetilde{\mathcal{A}}_{1122} & \widetilde{\mathcal{A}}_{1222} & \widetilde{\mathcal{A}}_{1322} & \widetilde{\mathcal{A}}_{1422} \\ \widetilde{\mathcal{A}}_{2121} & \widetilde{\mathcal{A}}_{2221} & \widetilde{\mathcal{A}}_{2321} & \widetilde{\mathcal{A}}_{2421} & \widetilde{\mathcal{A}}_{2122} & \widetilde{\mathcal{A}}_{2222} & \widetilde{\mathcal{A}}_{2322} & \widetilde{\mathcal{A}}_{2422} \\ \widetilde{\mathcal{A}}_{1131} & \widetilde{\mathcal{A}}_{1231} & \widetilde{\mathcal{A}}_{1331} & \widetilde{\mathcal{A}}_{1431} & \widetilde{\mathcal{A}}_{1132} & \widetilde{\mathcal{A}}_{1232} & \widetilde{\mathcal{A}}_{1332} & \widetilde{\mathcal{A}}_{1432} \\ \widetilde{\mathcal{A}}_{2131} & \widetilde{\mathcal{A}}_{2231} & \widetilde{\mathcal{A}}_{2331} & \widetilde{\mathcal{A}}_{2431} & \widetilde{\mathcal{A}}_{2132} & \widetilde{\mathcal{A}}_{2232} & \widetilde{\mathcal{A}}_{2332} & \widetilde{\mathcal{A}}_{2432} \end{array}\right] & \begin{array}{c} {\scriptstyle (1,1)} \\ {\scriptstyle (2,1)} \\ {\scriptstyle (1,2)} \\ {\scriptstyle (2,2)} \\ {\scriptstyle (1,3)} \\ {\scriptstyle (2,3)} \end{array} \\ \begin{array}{cccccccc} {\scriptstyle (1,1)} & {\scriptstyle (2,1)} & {\scriptstyle (3,1)} & {\scriptstyle (4,1)} & {\scriptstyle (1,2)} & {\scriptstyle (2,2)} & {\scriptstyle (3,2)} & {\scriptstyle (4,2)} \end{array} & \end{array}$$

where $\widetilde{\mathcal{A}}_{\alpha\beta\gamma\delta} = (\mathcal{A}_{\alpha\beta\gamma\delta})_{\mathbf{r}\times\mathbf{c}}$. Notice the multi-indexing of the block rows and columns.

3.4. A special case. Returning to the second example in section 2.4, suppose

$$\mathcal{B} \;=\; B^{(1)} \circ \cdots \circ B^{(d)},$$

where

$$B^{(\ell)} \in \mathbb{R}^{q_\ell \times n_\ell}$$

for $\ell = 1,\ldots,d$. Assume that $[\mathbf{u}^{(\ell)}, \mathbf{v}^{(\ell)}]$ is a blocking for $B^{(\ell)}$ and note that

$$\mathbf{M} \;=\; \left\{\mathbf{u}^{(1)}, \mathbf{v}^{(1)}, \ldots, \mathbf{u}^{(d)}, \mathbf{v}^{(d)}\right\} \tag{3.21}$$

is a blocking for $\mathcal{B}$. Let $B_{\mu,\tau}^{(\ell)}$ denote block (μ,τ) of $B^{(\ell)}$. If

$$\mathbf{k} \;=\; [\, i_1, j_1, \ldots, i_d, j_d],$$

then the $\mathbf{k}$th block of $\mathcal{B}$ is given by

$$\mathcal{B}_{\mathbf{k}} \;=\; B_{i_1,j_1}^{(1)} \circ \cdots \circ B_{i_d,j_d}^{(d)}.$$

If

$$\mathbf{r} = 1{:}2{:}2d,$$

$$\mathbf{c} = 2{:}2{:}2d,$$

$$\mathbf{R} = \left\{\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(d)}\right\}, \tag{3.22}$$

$$\mathbf{C} = \left\{\mathbf{v}^{(1)}, \ldots, \mathbf{v}^{(d)}\right\}, \tag{3.23}$$

then by applying (3.17) and (2.25) we see that

$$(\mathcal{B}_{\mathbf{R}\times\mathbf{C}})_{\mathbf{i},\mathbf{j}} \;=\; \left(B^{(1)}_{i_1,j_1} \circ \cdots \circ B^{(d)}_{i_d,j_d}\right)_{\mathbf{r}\times\mathbf{c}} \;=\; B^{(d)}_{i_d,j_d} \otimes \cdots \otimes B^{(1)}_{i_1,j_1}. \tag{3.24}$$

Here, the notation $(\mathcal{B}_{\mathbf{R}\times\mathbf{C}})_{\mathbf{i},\mathbf{j}}$ denotes block $(ivec(\mathbf{i},\mathbf{q}), ivec(\mathbf{j},\mathbf{n}))$. This result is key to the development of a block-level multilinear product which we pursue in section 4.2.

4. Blocked contractions.

We next apply our block tensor "technology" to the problem of computing a contraction between two tensors. A multi-index summation notation will be used to describe the summations. If $\mathbf{n}$ is a length-d index vector, then

$$\sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \equiv \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d}.$$

4.1. The general case.

It is instructive to work through a small, motivating example before we present the main results. Suppose we are given $\mathcal{F} \in \mathbb{R}^{\alpha_1\times\cdots\times\alpha_4}$ and $\mathcal{G} \in \mathbb{R}^{\beta_1\times\cdots\times\beta_5}$ and wish to compute the order-5 tensor $\mathcal{H} \in \mathbb{R}^{\alpha_3\times\alpha_4\times\beta_3\times\beta_4\times\beta_5}$ defined by

$$\mathcal{H}(i_1,i_2,j_1,j_2,j_3) \;=\; \sum_{k_1=1}^{\alpha_3}\sum_{k_2=1}^{\alpha_4} \mathcal{F}(i_1,i_2,k_1,k_2)\cdot\mathcal{G}(k_1,k_2,j_1,j_2,j_3). \tag{4.1}$$

Of course, for this to make sense, we must have $\alpha_3 = \beta_1$ and $\alpha_4 = \beta_2$. It is well known that a tensor contraction such as this can be "reshaped" into a single matrix-matrix multiplication. To see this we rewrite (4.1) using multi-index notation,

$$\mathcal{H}(\mathbf{i},\mathbf{j}) \;=\; \sum_{\mathbf{k}=\mathbf{1}}^{\boldsymbol{\alpha}(3:4)} \mathcal{F}(\mathbf{i},\mathbf{k})\cdot\mathcal{G}(\mathbf{k},\mathbf{j}). \tag{4.2}$$

Define the index vectors

$$\boldsymbol{r} \;=\; [\,1\;2\,], \qquad \boldsymbol{\lambda} \;=\; [\,3\;4\,], \qquad \boldsymbol{\psi} \;=\; [\,1\;2\,], \qquad \mathbf{c} \;=\; [\,3\;4\;5\,],$$

and note that $\mathbf{1} \le \mathbf{i} \le \boldsymbol{\alpha}(\mathbf{r})$ and $\mathbf{1} \le \mathbf{j} \le \boldsymbol{\beta}(\mathbf{c})$ in (4.2). Recall from (2.17)–(2.20) that the rows and columns of a tensor unfolding are vecs of reduced-order subtensors. In particular,

$$\mathcal{F}_{\mathbf{r}\times\boldsymbol{\lambda}}(\mathbf{i},:) = \mathrm{vec}(\mathcal{F}^{(\mathbf{i})})^T,$$

$$\mathcal{G}_{\boldsymbol{\psi}\times\mathbf{c}}(:,\mathbf{j}) = \mathrm{vec}(\mathcal{G}^{(\mathbf{j})}),$$

where $\mathcal{F}^{(\mathbf{i})} \in \mathbb{R}^{\alpha_3 \times \alpha_4}$ and $\mathcal{G}^{(\mathbf{j})} \in \mathbb{R}^{\beta_1 \times \beta_2}$ are defined by

$$\begin{array}{rclcrcl} \mathcal{F}^{(\mathbf{i})}(k_1,k_2) & = & \mathcal{F}(i_1,i_2,k_1,k_2), & & \mathbf{i} & = & [\, i_1\, i_2 \,], \\ \mathcal{G}^{(\mathbf{j})}(k_1,k_2) & = & \mathcal{G}(k_1,k_2,j_1,j_2,j_3), & & \mathbf{j} & = & [\, j_1\, j_2\, j_3 \,]. \end{array}$$

It follows from (4.2) that

$$\mathcal{H}(\mathbf{i},\mathbf{j}) \;=\; \sum_{k_1=1}^{\alpha_1}\sum_{k_2=1}^{\alpha_2} \mathcal{F}^{(\mathbf{i})}(k_1,k_2)\cdot \mathcal{G}^{(\mathbf{j})}(k_1,k_2) \;=\; \mathcal{F}_{\mathbf{r}\times\boldsymbol{\lambda}}(\mathbf{i},:)\cdot \mathcal{G}_{\boldsymbol{\psi}\times\mathbf{c}}(:,\mathbf{j}),$$

and thus

$$\mathcal{H}_{[1\,2]\times[3\,4\,5]} \;=\; \mathcal{F}_{\mathbf{r}\times\boldsymbol{\lambda}}\cdot \mathcal{G}_{\boldsymbol{\psi}\times\mathbf{c}}.$$

In this example, the summation is over the last two modes of $\mathcal{F}$ and the first two modes of $\mathcal{G}$. These are convenient locations for the summation indices because the contraction $\mathcal{H}$ is then easily seen to be "isomorphic" to a matrix-matrix product of simple tensor unfoldings.

If the summation modes are arbitrarily positioned, then they can be moved to these friendly locations through transposition. This result is widely known and exploited, e.g., in [2, 11]. Nevertheless, in keeping with the spirit of this paper we think that it is useful to include a formal verification of this important maneuver.

THEOREM 4.1. *Suppose* $\mathcal{F} \in \mathbb{R}^{\alpha_1\times\cdots\times\alpha_{f+\ell}}$, $\mathcal{G} \in \mathbb{R}^{\beta_1\times\cdots\times\beta_{g+\ell}}$, *and that* $\mathbf{p}$ *and* $\mathbf{q}$ *are permutations of* $1{:}f{+}\ell$ *and* $1{:}g{+}\ell$, *respectively. Define*

$$\begin{array}{rclcrcl} \mathbf{r} & = & \mathbf{p}(1{:}f), & & \boldsymbol{\lambda} & = & \mathbf{p}((f{+}1){:}(f{+}\ell)), \\ \boldsymbol{\psi} & = & \mathbf{q}(1{:}\,\ell), & & \mathbf{c} & = & \mathbf{q}((\,\ell{+}1){:}(\ell{+}g)), \end{array}$$

and assume $\boldsymbol{\alpha}(\boldsymbol{\lambda}) = \boldsymbol{\beta}(\boldsymbol{\psi})$. *If* $\mathcal{H} \in \mathbb{R}^{\alpha_{r_1}\times\cdots\times\alpha_{r_f}\times\beta_{c_1}\times\cdots\times\beta_{c_g}}$ *is defined by*

$$\mathcal{H}(\mathbf{i},\mathbf{j}) \;=\; \sum_{\mathbf{k}=\mathbf{1}}^{\boldsymbol{\alpha}(\boldsymbol{\lambda})} \mathcal{F}^{\langle\mathbf{p}\rangle}(\mathbf{i},\mathbf{k})\,\mathcal{G}^{\langle\mathbf{q}\rangle}(\mathbf{k},\mathbf{j}), \qquad \mathbf{1}\le\mathbf{i}\le\boldsymbol{\alpha}(\mathbf{r}), \quad \mathbf{1}\le\mathbf{j}\le\boldsymbol{\beta}(\mathbf{c}), \tag{4.3}$$

then

$$\mathcal{H}_{[1:f]\times[f+1:f+g]} \;=\; \mathcal{F}_{\mathbf{r}\times\boldsymbol{\lambda}}\cdot \mathcal{G}_{\boldsymbol{\psi}\times\mathbf{c}}. \tag{4.4}$$

Proof. The assumption $\boldsymbol{\alpha}(\boldsymbol{\lambda}) = \boldsymbol{\beta}(\boldsymbol{\psi})$ ensures that the summations in (4.3) are well defined. Using (2.17)–(2.20), we have

$$\mathcal{F}_{\mathbf{r}\times\boldsymbol{\lambda}}(\mathbf{i},:) = \mathrm{vec}(\mathcal{F}^{(\mathbf{i})})^T,$$
$$\mathcal{G}_{\boldsymbol{\psi}\times\mathbf{c}}(:,\mathbf{j}) = \mathrm{vec}(\mathcal{G}^{(\mathbf{j})}),$$

where $\mathcal{F}^{(\mathbf{i})} \in \mathbb{R}^{\alpha_{\lambda_1}\times\cdots\times\alpha_{\lambda_\ell}}$ and $\mathcal{G}^{(\mathbf{j})} \in \mathbb{R}^{\beta_{\psi_1}\times\cdots\times\beta_{\psi_\ell}}$ are defined by

$$\mathcal{F}^{(\mathbf{i})}(\mathbf{k}) = \mathcal{F}^{\langle\mathbf{p}\rangle}(i_1,\ldots,i_f,k_1,\ldots,k_\ell),$$
$$\mathcal{G}^{(\mathbf{j})}(\mathbf{k}) = \mathcal{G}^{\langle\mathbf{q}\rangle}(k_1,\ldots,k_\ell,j_1,\ldots,j_g).$$

It follows that for all $\mathbf{i}$ and $\mathbf{j}$ that satisfy $\mathbf{1} \le \mathbf{i} \le \boldsymbol{\alpha}(\mathbf{r})$ and $\mathbf{1} \le \mathbf{j} \le \boldsymbol{\beta}(\mathbf{c})$ we have

$$\begin{aligned}\mathcal{H}(\mathbf{i},\mathbf{j}) &= \sum_{\mathbf{k}=\mathbf{1}}^{\boldsymbol{\alpha}(\boldsymbol{\lambda})} \mathcal{F}^{\langle \mathbf{p}\rangle}(\mathbf{i},\mathbf{k}) \cdot \mathcal{G}^{\langle \mathbf{q}\rangle}(\mathbf{k},\mathbf{j}) \\ &= \sum_{\mathbf{k}=\mathbf{1}}^{\boldsymbol{\alpha}(\boldsymbol{\lambda})} \mathcal{F}^{(\mathbf{i})}(\mathbf{k}) \cdot \mathcal{G}^{(\mathbf{j})}(\mathbf{k}) \;=\; \mathcal{F}_{\mathbf{r}\times\boldsymbol{\lambda}}(\mathbf{i},:)\cdot \mathcal{G}_{\boldsymbol{\psi}\times\mathbf{c}}(:,\mathbf{j}),\end{aligned}$$

which, using (2.14)–(2.18), implies (4.4). ∎

It is instructive to illustrate what the theorem "says" when $\mathbf{c} = \emptyset$. Suppose $\mathcal{F} \in \mathbb{R}^{\alpha_1\times\cdots\times\alpha_5}$ and $\mathcal{G} \in \mathbb{R}^{\beta_1\times\beta_2}$ with $\alpha_2 = \beta_2$, $\alpha_3 = \beta_1$. If the tensor $\mathcal{H} \in \mathbb{R}^{\alpha_5\times\alpha_1\times\alpha_4}$ is defined by the contraction

$$\mathcal{H}(i_1, i_2, i_3) \;=\; \sum_{\mathbf{k}=\mathbf{1}}^{\boldsymbol{\alpha}(2:3)} \mathcal{F}(i_2, k_1, k_2, i_3, i_1)\mathcal{G}(k_2, k_1),$$

then in the notation of the theorem we have $f = 3, \ell = 2, g = 0$, $\mathbf{p} = [\,5\ 1\ 4\ 2\ 3\,]$, and $\mathbf{q} = [\,2\ 1\,]$. It follows that $\mathbf{r} = [\,5\ 1\ 4\,]$, $\mathbf{c} = \emptyset$, $\boldsymbol{\lambda} = [\,2\ 3\,]$, and $\boldsymbol{\psi} = [\,2\ 1\,]$. Thus, we may conclude from (4.4) that

$$\mathcal{H}_{[1:3]\times\emptyset} \;=\; \mathrm{vec}(\mathcal{H}) \;=\; \mathcal{F}_{[5\ 1\ 4]\times[2\ 3]} \cdot \mathcal{G}_{[2\ 1]\times\emptyset} \;=\; \mathcal{F}_{[5\ 1\ 4]\times[2\ 3]} \cdot \mathrm{vec}(\mathcal{G}^T),$$

a matrix-vector product.

If the tensors $\mathcal{F}$ and $\mathcal{G}$ are "blocked conformally," then (4.3) can be reformulated as a product of two block matrices.

Corollary 4.2. *Assume that the notation and conditions of Theorem* 4.1 *hold. Let*

$$\mathbf{S} \;=\; \{\mathbf{s}^{(1)}, \ldots, \mathbf{s}^{(f+\ell)}\} \tag{4.5}$$

be a blocking for $\mathcal{F}$*, and set*

$$\mathbf{R} \;=\; \{\mathbf{s}^{(r_1)}, \ldots, \mathbf{s}^{(r_f)}\}, \qquad \boldsymbol{\Lambda} \;=\; \{\mathbf{s}^{(\lambda_1)}, \ldots, \mathbf{s}^{(\lambda_\ell)}\}.$$

Likewise, let

$$\mathbf{T} \;=\; \{\mathbf{t}^{(1)}, \ldots, \mathbf{t}^{(g+\ell)}\} \tag{4.6}$$

be a blocking for $\mathcal{G}$*, and set*

$$\boldsymbol{\Psi} \;=\; \{\mathbf{t}^{(\psi_1)}, \ldots, \mathbf{t}^{(\psi_\ell)}\}, \qquad \mathbf{C} \;=\; \{\mathbf{t}^{(c_1)}, \ldots, \mathbf{t}^{(c_g)}\}.$$

If

$$\mathbf{s}^{(\lambda_k)} = \mathbf{t}^{(\psi_k)}, \qquad k = 1, \ldots, \ell, \tag{4.7}$$

then with respect to the tensor $\mathcal{H}$*,* $\mathbf{R}$ *is a blocking for modes* 1 *through* f*,* $\mathbf{C}$ *is a blocking for modes* $f+1$ *through* $f+g$*, and*

$$\mathcal{H}_{\mathbf{R}\times\mathbf{C}} \;=\; \mathcal{F}_{\mathbf{R}\times\boldsymbol{\Lambda}} \cdot \mathcal{G}_{\boldsymbol{\Psi}\times\mathbf{C}}. \tag{4.8}$$

Proof. From Theorem 3.3 we have

$$\mathcal{F}_{\mathbf{R}\times\mathbf{\Lambda}} \;=\; P_{\mathbf{R}}\mathcal{F}_{\mathbf{r}\times\boldsymbol{\lambda}}P_{\mathbf{\Lambda}}^{T}, \qquad\qquad \mathcal{G}_{\mathbf{\Psi}\times\mathbf{C}} \;=\; P_{\mathbf{\Psi}}\mathcal{G}_{\boldsymbol{\psi}\times\mathbf{c}}P_{\mathbf{C}}^{T}.$$

Since $\{\mathbf{s}^{(r_1)},\ldots,\mathbf{s}^{(r_f)},\mathbf{t}^{(c_1)},\ldots,\mathbf{t}^{(c_g)}\}$ is a blocking for $\mathcal{H}$, we also have

$$\mathcal{H}_{\mathbf{R}\times\mathbf{C}} \;=\; P_{\mathbf{R}}\cdot\mathcal{H}_{[1\,:\,f]\times[f+1\,:\,f+g]}\cdot P_{\mathbf{C}}^{T}.$$

The conformability condition (4.7) implies $P_{\mathbf{\Lambda}} = P_{\mathbf{\Psi}}$, and so it follows from (4.4) that

$$\begin{aligned}\mathcal{H}_{\mathbf{R}\times\mathbf{C}} &= P_{\mathbf{R}}(\mathcal{F}_{\mathbf{r}\times\boldsymbol{\lambda}}\cdot\mathcal{G}_{\boldsymbol{\psi}\times\mathbf{c}})P_{\mathbf{C}}^{T}\\ &= (P_{\mathbf{R}}\mathcal{F}_{\mathbf{r}\times\boldsymbol{\lambda}}P_{\mathbf{\Lambda}}^{T})(P_{\mathbf{\Psi}}\mathcal{G}_{\boldsymbol{\psi}\times\mathbf{c}}P_{\mathbf{C}}^{T}) \;=\; \mathcal{F}_{\mathbf{R}\times\mathbf{\Lambda}}\cdot\mathcal{G}_{\mathbf{\Psi}\times\mathbf{C}},\end{aligned}$$

completing the proof. ∎

Thus, the tensor $\mathcal{H}$ in (4.3) can be computed as either a matrix product (4.4) or a block matrix product (4.8). For the latter case, we develop recipes for the blocks of $\mathcal{H}_{\mathbf{R}\times\mathbf{C}}$. Let $b_j^{(\mathrm{S})}$ be the length of the blocking vector $\mathbf{s}^{(j)}$ in (4.5), and let $b_j^{(\mathrm{T})}$ be the length of the blocking vector $\mathbf{t}^{(j)}$ in (4.6). Note that if

$$\begin{aligned} b_{rows}^{(\mathcal{F})} &= b_{r_1}^{(\mathrm{S})}\cdots b_{r_f}^{(\mathrm{S})}, & b_{cols}^{(\mathcal{F})} &= b_{\lambda_1}^{(\mathrm{S})}\cdots b_{\lambda_\ell}^{(\mathrm{S})},\\ b_{rows}^{(\mathcal{G})} &= b_{\psi_1}^{(\mathrm{T})}\cdots b_{\psi_\ell}^{(\mathrm{T})}, & b_{cols}^{(\mathcal{G})} &= b_{c_1}^{(\mathrm{T})}\cdots b_{c_g}^{(\mathrm{T})},\end{aligned}$$

then (4.7) implies $b_{cols}^{(\mathcal{F})} = b_{rows}^{(\mathcal{G})}$, and we observe that

$$\left\{\begin{array}{c}\mathcal{F}_{\mathbf{R}\times\mathbf{\Lambda}}\\[4pt] \mathcal{G}_{\mathbf{\Lambda}\times\mathbf{C}}\\[4pt] \mathcal{H}_{\mathbf{R}\times\mathbf{C}}\end{array}\right\} \quad\text{is a}\quad \left\{\begin{array}{c} b_{rows}^{(\mathcal{F})}\text{-by-}b_{cols}^{(\mathcal{F})}\\[4pt] b_{rows}^{(\mathcal{G})}\text{-by-}b_{cols}^{(\mathcal{G})}\\[4pt] b_{rows}^{(\mathcal{F})}\text{-by-}b_{cols}^{(\mathcal{G})}\end{array}\right\}\quad\text{block matrix.}$$

If $\mathbf{1} \le \boldsymbol{\mu} \le \mathbf{b}^{(\mathrm{S})}(\mathbf{r})$ and $\mathbf{1} \le \boldsymbol{\tau} \le \mathbf{b}^{(\mathrm{T})}(\mathbf{c})$, and $\mu = ivec(\boldsymbol{\mu},\mathbf{b}^{(\mathrm{S})}(\mathbf{r}))$ and $\tau = ivec(\boldsymbol{\tau},\mathbf{b}^{(\mathrm{T})}(\mathbf{c}))$, then block (μ,τ) of $\mathcal{H}_{\mathbf{R}\times\mathbf{C}}$ is given by

$$(\mathcal{H}_{\mathbf{R}\times\mathbf{C}})_{\boldsymbol{\mu},\boldsymbol{\tau}} \;=\; \sum_{\mathbf{q}=\mathbf{1}}^{\mathbf{b}^{(\mathrm{S})}(\boldsymbol{\lambda})} (\mathcal{F}_{\mathbf{R}\times\mathbf{\Lambda}})_{\boldsymbol{\mu},\mathbf{q}}\,(\mathcal{G}_{\mathbf{\Psi}\times\mathbf{C}})_{\mathbf{q},\boldsymbol{\tau}}.$$

Using (3.17) this can be rewritten in terms of subtensor unfoldings. Indeed, if index vectors $\mathbf{k}$, $\mathbf{i}^{(\mathbf{q})}$, and $\mathbf{j}^{(\mathbf{q})}$ are defined by

$$\begin{aligned} \mathbf{k}(\mathbf{r}) &= \boldsymbol{\mu}, & \mathbf{k}(\mathbf{c}) &= \boldsymbol{\tau},\\ \mathbf{i}^{(\mathbf{q})}(\mathbf{r}) &= \mathbf{k}(\mathbf{r}), & \mathbf{i}^{(\mathbf{q})}(\boldsymbol{\lambda}) &= \mathbf{q},\\ \mathbf{j}^{(\mathbf{q})}(\boldsymbol{\psi}) &= \mathbf{q}, & \mathbf{j}^{(\mathbf{q})}(\mathbf{c}) &= \mathbf{k}(\mathbf{c}),\end{aligned}$$

then

$$(\mathcal{H}_{\mathbf{k}})_{[1:f]\times[f+1:f+g]} \;=\; \sum_{\mathbf{q}=\mathbf{1}}^{\mathbf{b}^{(\mathrm{S})}(\boldsymbol{\lambda})} (\mathcal{F}_{\mathbf{i}^{(\mathbf{q})}})_{\mathbf{r}\times\boldsymbol{\lambda}}\,(\mathcal{G}_{\mathbf{j}^{(\mathbf{q})}})_{\boldsymbol{\psi}\times\mathbf{c}}. \tag{4.9}$$

4.2. Blocked multilinear products. As an example of how the preceding results can be adapted to handle structured contractions, we briefly consider the multilinear product since we have developed the supporting formulae in sections 2.4 and 3.4. Suppose $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ and that

$$B^{(k)} \in \mathbb{R}^{q_k \times n_k}, \qquad k = 1, \ldots, d.$$

The tensor $\mathcal{C} \in \mathbb{R}^{q_1 \times \cdots \times q_d}$ specified by

$$\mathcal{C}(\mathbf{i}) \;=\; \sum_{\mathbf{k}=\mathbf{1}}^{\mathbf{n}} \mathcal{A}(\mathbf{k}) B^{(1)}(i_1, k_1) \cdots B^{(d)}(i_d, k_d) \tag{4.10}$$

is the *multilinear product* of $\mathcal{A}$ with $B^{(1)}, \ldots, B^{(d)}$ and is denoted [5] by

$$\mathcal{C} \;=\; (B^{(1)}, \ldots, B^{(d)}) \cdot \mathcal{A}.$$

If the order-$(2d)$ tensor $\mathcal{B}$ is defined by

$$\mathcal{B} \;=\; B^{(1)} \circ \cdots \circ B^{(d)},$$

then we see that $\mathcal{C}$ is a contraction of the form

$$\mathcal{C}(\mathbf{i}) \;=\; \sum_{\mathbf{k}=\mathbf{1}}^{\mathbf{n}} \mathcal{A}(\mathbf{k}) \mathcal{B}(i_1, k_1, \ldots, i_d, k_d).$$

We apply Theorem 4.1 with $\mathcal{F} = \mathcal{B}$, $f = d$, $\ell = d$, $\mathcal{G} = \mathcal{A}$, $g = 0$, $\mathbf{r} = 1{:}2{:}2d$, $\boldsymbol{\lambda} = 2{:}2{:}2d$, $\boldsymbol{\psi} = 1{:}d$, and $\mathbf{c} = \emptyset$. It follows that $\mathcal{A}_{\boldsymbol{\psi} \times \mathbf{c}} = \mathrm{vec}(\mathcal{A})$ and $\mathcal{C}_{[1:\ell] \times [\ell+1:\ell]} = \mathrm{vec}(\mathcal{C})$, and so from Theorem 4.1 and (2.25) we have

$$\mathrm{vec}(\mathcal{C}) \;=\; \left(B^{(d)} \otimes \cdots \otimes B^{(1)}\right) \mathrm{vec}(\mathcal{A}). \tag{4.11}$$

If the B matrices are blocked according to (3.21) and $\mathbf{R}$ and $\mathbf{C}$ are defined by (3.22)–(3.23), then $\mathbf{R}$ is a blocking for $\mathcal{C}$, $\mathbf{C}$ is a blocking for $\mathcal{A}$, and

$$P_{\mathbf{R}} \mathrm{vec}(\mathcal{C}) \;=\; \left(P_{\mathbf{R}} \left(B^{(d)} \otimes \cdots \otimes B^{(1)}\right) P_{\mathbf{C}}^T\right) P_{\mathbf{C}} \mathrm{vec}(\mathcal{A}). \tag{4.12}$$

From (3.24) we see that the matrix

$$\mathcal{B}_{\mathbf{R} \times \mathbf{C}} \;=\; P_{\mathbf{R}} \left(B^{(d)} \otimes \cdots \otimes B^{(1)}\right) P_{\mathbf{C}}^T \tag{4.13}$$

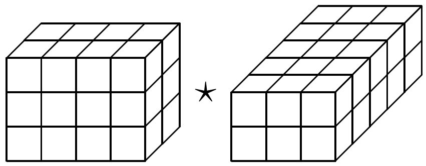
is a block matrix whose entries are Kronecker products. Indeed, $\mathcal{B}_{\mathbf{R} \times \mathbf{C}}$ is essentially the Tracy–Singh product of the B matrices; see [16]. Thus, from (4.11)–(4.13) we have the following block specification for $\mathcal{C}$:

$$\mathrm{vec}_{\mathbf{R}}(\mathcal{C}) \;=\; \mathcal{B}_{\mathbf{R} \times \mathbf{C}} \mathrm{vec}_{\mathbf{C}}(\mathcal{A}). \tag{4.14}$$

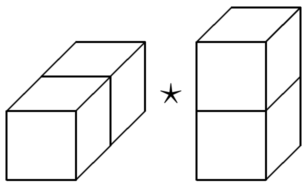
4.3. Visualization. As in block matrix computations, it is sometimes important to view a given blocked tensor contraction from different viewpoints. A small example builds an appreciation for this point.

Suppose $\mathcal{F}$ is a $3 \times 4 \times 2$ block tensor and $\mathcal{G}$ is a $2 \times 3 \times 5$ block tensor such that the blockings in mode 3 in $\mathcal{F}$ and mode 1 in $\mathcal{G}$ conform. Let $\mathcal{H}$ be the $3 \times 4 \times 3 \times 5$ block tensor whose elements are given by

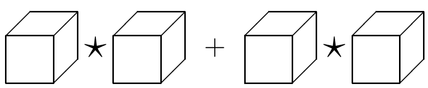
$$\mathcal{H}(i_1, i_2, j_1, j_2) = \sum_k \mathcal{F}(i_1, i_2, k) \cdot \mathcal{G}(k, j_1, j_2).$$



(1) The tensor contraction $\mathcal{H} = \mathcal{F} \star \mathcal{G}$ of two order-3 tensors viewed graphically as a contraction of conformally blocked tensors.



(2) Block $\mathcal{H}_{abcd} = \mathcal{H}(\alpha_1{:}\alpha_2, \beta_1{:}\beta_2, \gamma_1{:}\gamma_2, \delta_1{:}\delta_2)$ is a $\star$-contraction of two "block fibers," one from $\mathcal{F}$ and one from $\mathcal{G}$, i.e., $\mathcal{H}_{abcd} = \mathcal{F}(\alpha_1{:}\alpha_2, \beta_1{:}\beta_2, :) \star \mathcal{G}(:, \gamma_1{:}\gamma_2, \delta_1{:}\delta_2)$.



(3) The $\star$-contraction of the two block fibers is a sum of $\star$-contractions of fiber blocks, i.e., $\mathcal{H}_{abcd} = \mathcal{F}_{ab1} \star \mathcal{G}_{1cd} + \mathcal{F}_{ab2} \star \mathcal{G}_{2cd}$.

FIG. 4.1. *Three levels of a blocked contraction.*

For convenience, denote the operation of contracting two order-3 tensors $\mathcal{T}_1$ and $\mathcal{T}_2$ in this way as $\mathcal{T}_1 \star \mathcal{T}_2$, e.g., $\mathcal{H} = \mathcal{F} \star \mathcal{G}$. Figure 4.1 shows how this blocked contraction can be visualized at three different levels. At the lowest level, block $[a, b, c, d]$ in $\mathcal{H}$ can be computed via the matrix equation

$$(\mathcal{H}_{abcd})_{[1\ 2]\times[3\ 4]} = (\mathcal{F}_{ab1})_{[1\ 2]\times[3]} \cdot (\mathcal{G}_{1cd})_{[1]\times[2\ 3]} + (\mathcal{F}_{ab2})_{[1\ 2]\times[3]} \cdot (\mathcal{G}_{2cd})_{[1]\times[2\ 3]}.$$

This follows from (4.9) and is depicted in part (3) of Figure 4.1.

5. Concluding remarks. Given the nature of this paper, it is important to be reminded in this closing section that there is a big difference between a cryptic mathematical formula and its utilization in practice. A case in point is the permutation matrix $P_{\mathbf{M}}$ that is characterized in Theorem 3.1. Obviously, an integer vector should be used to represent a permutation matrix like $P_{\mathbf{M}}$; it should never be computed as a two-dimensional array. We offer a few details based on the convention that if $P = I_n(:, \mathbf{v})$ where $\mathbf{v}$ is permutation of $1{:}n$, then $\mathbf{v}$ represents P. We capture this connection with the notation $P_{\mathbf{v}}$. Note that if $y = P_{\mathbf{v}} x$, then $y = x(\mathbf{v})$, while $y(\mathbf{v}) = x$ implies $y = P_{\mathbf{v}}^T x$. Letting $\mathbf{1}_n$ denote the n-vector of ones, here are some basic facts that concern this style of representation:

1. If q and r are positive integers and $\mathbf{w} = [1{:}r{:}qr \;\; 2{:}2{:}qr \;\; \cdots \;\; r{:}r{:}qr]$, then $P_{\mathbf{w}} = \Pi_{q,r}$, the (q, r) perfect shuffle.
2. If $\mathbf{u}$ and $\mathbf{v}$ are permutations of $1{:}n$ and $\mathbf{w} = \mathbf{v}(\mathbf{u})$, then $P_{\mathbf{w}} = P_{\mathbf{u}} P_{\mathbf{v}}$.
3. If $\mathbf{u}$ is a permutation of $1{:}n$ and $\mathbf{v}$ is a permutation of $1{:}m$, then $P_{\mathbf{w}} = P_{\mathbf{u}} \otimes P_{\mathbf{v}}$, where $\mathbf{w} = \mathbf{1}_n \otimes \mathbf{v} \, + \, m \cdot (\mathbf{u} - \mathbf{1}_n) \otimes \mathbf{1}_m$.
4. If $\mathbf{u}$ is a permutation of $1{:}n$ and $\mathbf{v}$ is a permutation of $1{:}m$, then $P_{\mathbf{w}} = \mathrm{diag}(P_{\mathbf{u}}, P_{\mathbf{v}})$, where $\mathbf{w} = [\mathbf{u} \;\; (n \cdot \mathbf{1}_n + \mathbf{v})]$.

The vector representation of the matrix $P_{\mathbf{M}}$, since it is defined by perfect shuffles, Kronecker products, and direct sums, can be efficiently assembled using these facts.

Another illustration of the gap between formula and implementation concerns (4.11). The calculation of a multilinear product $\mathcal{C} = (B^{(1)}, \ldots, B^{(d)}) \cdot \mathcal{A}$ would not explicitly use this formula. Instead it would proceed as follows:

for $i = 1, \ldots, d$

$$\mathcal{A} \leftarrow (I_{n_1}, \ldots, B^{(i)}, \ldots, I_{n_d}) \cdot \mathcal{A}$$

end

The ith update is referred to as the *i-mode product*; see [4, 13]. By using Theorem 4.1 we see that this is equivalent to the matrix-matrix multiplication

$$\mathcal{A}_{(i)} \leftarrow B^{(i)} \mathcal{A}_{(i)},$$

where $\mathcal{A}_{(i)} \equiv \mathcal{A}_{[i] \times [1:i-1\ i+1:d]}$ is the mode-i unfolding of $\mathcal{A}$ mentioned in section 2.3.

Similarly, in a block-based implementation of the multilinear product, one would not directly use (4.14). Instead, the block-matrix multiplications

$$\mathcal{A}_{\mathbf{J} \times \mathbf{C}} \leftarrow B^{(i)} \mathcal{A}_{\mathbf{I} \times \mathbf{C}}$$

would be carried out sequentially for modes $i = 1, \ldots, d$. Here, $\mathbf{I}$ is the original blocking for mode i, $\mathbf{J}$ is the new blocking of mode i inherited from the row blocking of $B^{(i)}$, and $\mathbf{C}$ is a blocking for modes $[1:i-1\ i+1:d]$ of $\mathcal{A}$.

Overall, it is reasonable to conclude from the above that block tensors behave in much the same way as block matrices. Although the precise formulas are more involved, the basic intuition that “all operations can be done at the block level” is correct. By making precise the notion of a block unfolding and developing a framework for reasoning about block tensor computation, we hope that we have laid a modest foundation for further research. Our own agenda includes looking at block versions of the tensor contraction engine [2], developing recursive tensor data structures that extend the clever ideas in [6], expanding the functionality of the Tensor Toolbox [11, 12] so that it supports block tensor computation, and analyzing block versions of various tensor iterations such as [3]. Throughout all this it will be important to chip away at the “notational divide” that currently besets the tensor computation community; see [9].

Acknowledgment. The authors are indebted to the two referees whose suggestions and corrections led to major improvements over the first and second drafts of the manuscript.

REFERENCES

[1] G. Ballard, J. Demmel, O. Holtz, and O. Schwartz, *Minimizing Communication in Numerical Linear Algebra*, Tech. rep. UCB/EECS-2009-62, University of California at Berkeley, Berkeley, CA, 2010.

[2] G. Baumgartner, A. Auer, D. Bernholdt, A. Bibireata, V. Choppella, D. Cociorva, X. Gao, R. Harrison, S. Hirata, S. Krishnamoorthy, S. Krishnan, C. Lam, Q. Lu, M. Nooijen, R. Pitzer, J. Ramanujam, P. Sadayappan, and A. Sibiryakov, *Synthesis of high-performance parallel programs for a class of ab initio quantum chemistry models*, Proc. IEEE, 93 (2005), pp. 276–292.

[3] L. De Lathauwer, P. Comon, B. De Moor, and J. Vandewalle, *Higher-order power method–application in independent component analysis*, in Proceedings of the International Symposium on Nonlinear Theory and Its Applications (NOLTA ’95), Las Vegas, NV, 1995, pp. 91–96.

[4] L. De Lathauwer, B. De Moor, and J. Vandewalle, *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1253–1278.

[5] V. De Silva and L.-H. Lim, *Tensor rank and the ill-posedness of the best low-rank approximation problem*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1084–1127.

[6] E. Elmroth, F. Gustavson, I. Jonsson, and B. Kågström, *Recursive blocked algorithms and hybrid data structures for dense matrix library software*, SIAM Rev., 46 (2004), pp. 3–45.

[7] H. V. Henderson and S. R. Searle, *The vec-permutation matrix, the vec operator, and Kronecker products, a review*, Linear Multilinear Algebra, 9 (1981), pp. 271–288.

[8] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.

[9] H A. L. Kiers, *Towards a standardized notation and terminology in multiway analysis*, J. Chemometr., 14 (2000), pp. 105–122.

[10] T. G. Kolda, *Multilinear Operators for Higher-Order Decompositions*, Tech. rep. SAND2006-2081, Sandia National Laboratories, Albuquerque, NM, and Livermore, CA, 2006.

[11] T. G. Kolda and B. W. Bader, *Algorithm* 862: *MATLAB tensor classes for fast algorithm prototyping*, ACM Trans. Math. Software, 32 (2006), pp. 635–653.

[12] T. G. Kolda and B. W. Bader, Matlab *Tensor Toolbox Version* 2.4, 2006, available online from http://csmr.ca.sandia.gov/∼tkolda/TensorToolbox/, 2010.

[13] T. G. Kolda and B. W. Bader, *Tensor decompositions and applications*, SIAM Rev., 51 (2009), pp. 455–500.

[14] A. H. Phan and A. Cichocki, *Block decomposition for very large-scale nonnegative tensor factorization*, in Proceedings of the 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Aruba, 2009, pp. 316–319.

[15] S. Ragnarsson and C. F. Van Loan, *Block tensors and symmetric embeddings*, Linear Algebra Appl., to appear; also available online from http://arxiv.org/abs/1010.0707.

[16] D. S. Tracy and R. P. Singh, *A new matrix product and its applications in partitioned matrices*, Statist. Neerlandica, 26 (1972), pp. 143–157.

[17] C. F. Van Loan, *Computational Frameworks for the Fast Fourier Transform*, SIAM, Philadelphia, 1992.

[18] C. F. Van Loan, *The ubiquitous Kronecker product*, J. Comput. Appl. Math., 123 (1999), pp. 85–100.