SCIENCE CHINA
Mathematics

• ARTICLES •

January 2015 Vol. 58 No. 1: 197–212
doi: 10.1007/s11425-014-4930-z

Linear operators and positive semidefiniteness of symmetric tensor spaces

LUO ZiYan[1,*], QI LiQun[2] & YE YinYu[3]

[1]*State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing* 100044*, China;*
[2]*Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China;*
[3]*Department of Management Science and Engineering, Stanford University, Stanford, CA* 94305*, USA*
Email: zyluo@bjtu.edu.cn, maqilq@polyu.edu.hk, yinyu-ye@stanford.edu

Received August 26, 2013; accepted March 2, 2014; published online November 14, 2014

Abstract We study symmetric tensor spaces and cones arising from polynomial optimization and physical sciences. We prove a decomposition invariance theorem for linear operators over the symmetric tensor space, which leads to several other interesting properties in symmetric tensor spaces. We then consider the positive semidefiniteness of linear operators which deduces the convexity of the Frobenius norm function of a symmetric tensor. Furthermore, we characterize the symmetric positive semidefinite tensor (SDT) cone by employing the properties of linear operators, design some face structures of its dual cone, and analyze its relationship to many other tensor cones. In particular, we show that the cone is self-dual if and only if the polynomial is quadratic, give specific characterizations of tensors that are in the primal cone but not in the dual for higher order cases, and develop a complete relationship map among the tensor cones appeared in the literature.

Keywords symmetric tensor, symmetric positive semidefinite tensor cone, linear operator, SOS cone

MSC(2010) 15A69, 53A45, 47A05, 53C35

Citation: Luo Z Y, Qi L Q, Ye Y Y. Linear operators and positive semidefiniteness of symmetric tensor spaces. Sci China Math, 2015, 58: 197–212, doi: 10.1007/s11425-014-4930-z

1 Introduction

A symmetric tensor is a higher order generalization of a symmetric matrix. Analogous to the fact that a symmetric matrix can be regarded both as a linear operator on some Euclidean space and as an element in symmetric matrices space, a symmetric tensor both acts as a multilinear operator on some Cartesian products of Euclidean spaces and as a component of symmetric tensors space. As the former one, it has recently gained intense attention due to its wide applications to polynomial optimization [9], higher order derivatives of smooth functions [13], moments and cumulants of random vectors [15] in the theoretical respect, and to physical sciences such as imaging technologies in the practical respect. As the latter one, new mathematical developments for symmetric tensors involve tensor eigenvalues [17], tensor ranks and symmetric outer product decompositions [6], all of which are generalizations to symmetric matrices. Linear operators are fundamental and essential for any linear space, and hence for the symmetric tensor space. However, few papers concentrate on linear operators and their properties which work on symmetric tensors in the manner of a tensor-level thinking.

*Corresponding author

In this paper, we are interested in a class of special linear operators on symmetric tensor spaces, which further contribute to exploiting properties of some special components — the cone of positive semidefinite symmetric tensors (SDT cone for short) in symmetric tensor spaces. An intimate relation to positive semidefinite symmetric tensors is the nonnegative homogeneous polynomials, where the nonnegativity is an intrinsic property of polynomial functions, as one can see from quadratic polynomial functions. Let $x \in \mathbb{R}^n$ and m be a positive integer. An m-order homogeneous polynomial function in x can be written as

$$f_A(x) = \sum_{i_1,\dots,i_m=1}^{n} A_{i_1\cdots i_m} x_{i_1} x_{i_2} \cdots x_{i_m}. \tag{1.1}$$

The coefficients of the polynomial can be regarded as an m-order n-dimensional real tensor, denoted by A, which is symmetric — invariant under any permutation of its indices. Such a tensor A is said to be *positive semidefinite* if the corresponding polynomial $f_A(x)$ is nonnegative for any $x \in \mathbb{R}^n$. Obviously, positive semidefinite tensors only occur when m is even. The SDT cone consists of all such tensors. In this regard, verifying the nonnegativity of a homogeneous polynomial is equivalent to identifying the associated tensor lies in the SDT cone. Additionally, for many real applications, a tensor, either second or higher order, must be positive semidefinite to be physically meaningful. For example, the involved diffusion tensor in the diffusion weighted MRI is a symmetric positive semidefinite tensor of order 4 and dimension 3 (see [3]). All these show the evidence that to study the SDT cone is a necessary and meaningful work to a certain extent.

We greatly employ a special class of linear operators to get a better understanding of the SDT cone and the corresponding tensor space. Such a class of special linear operators on symmetric tensor spaces is just an m-order generalization of the symmetric-scaled representation used in interior-point algorithms in the literature of semidefinite programming [21], and shares a decomposition invariance theorem resembling that of the symmetric-scaled representation. Several other interesting properties in symmetric tensor spaces follow from this invariance property. All these properties may be potential to build a foundation for us to learn more about the symmetric tensor structures which provide greater descriptive flexibility than that of the corresponding reshaped matrix-based structures.

The positive semidefiniteness (i.e., monotonicity and self-adjointness) of such a class of linear operators is also considered which leads to the convexity of the Frobenius norm function of a symmetric tensor. Specially, for the case $n = 3$, we exhibit the matrix representation and the rank of such a matrix of the aforementioned class of linear operators, and then exploit the positive semidefiniteness of linear operators by the property of the corresponding matrices. Here the concept of matrix representation for a general linear operator was first proposed by Qi and Ye [18] by employing a linear isomorphic operator $\mathcal{L}$ which builds up a one-to-one correspondence from a tensor of dimension 3 to a real vector in $\mathbb{R}^\kappa$ with $\kappa = \frac{(m+2)(m+1)}{2}$, the number of independent components of a 3-dimensional symmetric tensor.

Based on the achieved properties of such a class of linear operators aforementioned, we proceed with characterizing the SDT cone. Just like all positive semidefinite symmetric matrices form a closed pointed convex self-dual cone in symmetric matrices space, the SDT cone can be easily verified as a closed pointed convex cone, while generally no longer self-dual in symmetric tensors space. In particular, we show that the symmetric positive semidefinite tensor cone is self-dual if and only if $m = 2$. We also present other results for the dual of the SDT cone such as a class of its faces, and give specific characterizations of tensors that in the primal cone but not in the dual. In addition, we analyze SDT's relationship to many other tensor cones, such as nonnegative tensor, partially diagonal tensor, and sum-of-squares (SOS) cones. We present a complete relationship map among these cones. This also expresses some striking differences between tensors and matrices.

The organization of this paper is as follows. In Section 2, we propose a decomposition invariance property for a class of linear operators and describe some other interesting properties of such linear operators based on such decomposition invariance property. In Section 3, the positive semidefiniteness of such linear operators is considered. Symmetric positive semidefinite tensor cone and its relationship to its dual cone and several other cones, such as nonnegative tensor cone, partially diagonal tensor cone and sum-of-squares (SOS) cone are characterized by the aforementioned properties of linear operators in Sections 4 and 5. Some concluding remarks are made in Section 6.

Before we proceed, we describe some notation used throughout this paper: $\mathcal{T}(m,n)$ stands for the set of all m-order n-dimensional real symmetric tensors; $[x]^m := x\otimes x\otimes\cdots\otimes x$ is a rank-one symmetric tensor generated by the outer product of m copies of a vector x; $\mathcal{S}(m,n)$ denotes the set of all positive semidefinite tensors in $\mathcal{T}(m,n)$; $\mathcal{V}(m,n)$ is the dual cone of $\mathcal{S}(m,n)$ in $\mathcal{T}(m,n)$; $\mathcal{L}(\cdot)$ is the linear isomorphic operator between $\mathcal{T}(m,3)$ and $\mathbb{R}^\kappa$ with $\kappa = \frac{1}{2}(m+1)(m+2)$. Let F be a convex subset of some convex closed cone K, $F \trianglelefteq K$ means that F is a face of K, i.e., for any x, $y \in K$ with $x+y \in F$, we have x, $y \in F$. The inner product of two tensors X, $Y \in \mathcal{T}(m,n)$ is defined as the component-wise product $\langle X, Y\rangle = \sum_{i_1,\ldots,i_m=1}^n X_{i_1\cdots i_m} Y_{i_1\cdots i_m}$ and $\|\cdot\|_F$ denotes the Frobenius norm which is deduced by the inner product. Unless otherwise pointed out, we will restrict m to be even in the rest of the paper.

2 A decomposition invariance property

A linear operator is a basic operation to any linear space, so is to the tensor space. Qi and Ye [18] proposed a class of linear operators based on matrices in tensors space for the purpose of the orthogonal similarity of E-eigenvalues. Comon and Sorensen [7] employed such class of linear operators based on orthogonal matrices for tensor diagonalization. The definition of this class of linear operators is reviewed as follows.

Definition 2.1. Let $P = (P_{ij}) \in \mathbb{R}^{n\times n}$. Define a linear operator $P^{[m]}$ such that

$$(P^{[m]}(A))_{i_1\cdots i_m} := \sum_{i'_1,\ldots,i'_m=1}^{n} P_{i_1 i'_1}\cdots P_{i_m i'_m} A_{i'_1\cdots i'_m}, \tag{2.1}$$

for any $A \in \mathcal{T}(m,n)$ with the $(i_1,\ldots,i_m)$-th entry $A_{i_1\cdots i_m}$.

We use $P^{[m]}$ to denote such a linear operator in this paper instead of P^m used in [17] to avoid the notation confusion from power operation of matrices in subsequent analyses. Note that $P^{[m]}$ is a linear operator from $\mathcal{T}(m,n)$ to itself according to [17, Proposition 10], and its definition is given in an entry-wise manner. Now, we turn our attention to the decomposition form of such linear operator based on the so-called symmetric outer product decomposition of symmetric tensors; see Comon et al. [6], i.e., for any $A \in \mathcal{T}(m,n)$, one can find $\alpha_1,\ldots,\alpha_s \in \mathbb{R}$ and $u^{(1)},\ldots,u^{(s)} \in \mathbb{R}^n$ such that

$$A = \sum_{i=1}^{s} \alpha_i [u^{(i)}]^m. \tag{2.2}$$

Though the symmetric outer product decomposition (2.2) of any given A is not unique in general [6,12], we prove that the linear operator $P^{[m]}$ enjoys the following decomposition invariance property.

Theorem 2.2 (Decomposition invariance). *For any $P \in \mathbb{R}^{n\times n}$, and any $A \in \mathcal{T}(m,n)$ with any of its symmetric outer product decomposition $A = \sum_{i=1}^s \alpha_i [u^{(i)}]^m$, we have*

$$P^{[m]}(A) = \sum_{i=1}^{s} \alpha_i [P(u^{(i)})]^m. \tag{2.3}$$

Proof. For any $A \in \mathcal{T}(m,n)$ and any $i_1,\ldots,i_m \in \{1,2,\ldots,n\}$, we have

$$\begin{aligned}
\left(\sum_{i=1}^{s} \alpha_i [P(u^{(i)})]^m\right)_{i_1\cdots i_m} &= \sum_{i=1}^{s} \alpha_i (P(u^{(i)}))_{i_1}\cdots (P(u^{(i)}))_{i_m} \\
&= \sum_{i=1}^{s} \alpha_i \left(\sum_{j_1=1}^{n} P_{i_1 j_1} u^{(i)}_{j_1}\right)\cdots\left(\sum_{j_m=1}^{n} P_{i_m j_m} u^{(i)}_{j_m}\right) \\
&= \sum_{i=1}^{s} \alpha_i \sum_{j_1,\ldots,j_m=1}^{n} P_{i_1 j_1} u^{(i)}_{j_1}\cdots P_{i_m j_m} u^{(i)}_{j_m}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j_1,\ldots,j_m=1}^{n} P_{i_1j_1}\cdots P_{i_mj_m} \sum_{i=1}^{s} \alpha_i u_{j_1}^{(i)} \cdots u_{j_m}^{(i)} \\
&= \sum_{j_1,\ldots,j_m=1}^{n} P_{i_1j_1}\cdots P_{i_mj_m} A_{j_1\cdots j_m} \\
&= (P^{[m]}(A))_{i_1\cdots i_m},
\end{aligned}
$$

where the last equality is achieved by (2.1). This completes the proof. □

It is worth pointing out that $P^{[m]}$ can be regarded as an extension of the symmetric-scaled representation Q_P in the setting of symmetric matrices space, i.e.,

$$
Q_P(A) = PAP^{\mathrm{T}} = P^{[2]}(A), \quad \forall\, A \in \mathcal{S}(2,n).
$$

Such an operator plays an essential role in the design of interior-point methods for the symmetry preservation in the context of semidefinite programming [2, 21] and has also been generalized to the Euclidean Jordan algebra, which is called the quadratic representation; see [8, 19, 20] for more details.

Next, we show that the operator $P^{[m]}$ possesses several nice properties analogous with the ones of quadratic representation. The decomposition invariance property developed in Theorem 2.2 greatly facilitates the proofs of these properties.

Proposition 2.3. *For any $P \in \mathbb{R}^{n\times n}$, we have*

(i) $(P^{[m]})^k = (P^k)^{[m]}$;

(ii) *if P is invertible, then* $(P^{[m]})^{-1} = (P^{-1})^{[m]}$;

(iii) $\langle P^{[m]}(A), B\rangle = \langle A, (P^{\mathrm{T}})^{[m]}(B)\rangle$, *for any A, $B \in \mathcal{T}(m,n)$*;

(iv) *if P is invertible, then* $\langle P^{[m]}(A), (P^{-T})^{[m]}(B)\rangle = \langle A, B\rangle$, *for any A, $B \in \mathcal{T}(m,n)$*;

(v) *if m is odd, then $P^{[m]}$ is a symmetric operator iff P is a symmetric matrix*;

(vi) *if m is even, then $P^{[m]}$ is a symmetric operator iff P is a symmetric matrix or a skew-symmetric matrix.*

Proof. The assertions in (i) and (ii) follow directly from the decomposition invariance property of $P^{[m]}$. For any A, $B \in \mathcal{T}(m,n)$, with $A = \sum_{i=1}^{s} \alpha_i [u^{(i)}]^m$, $B = \sum_{j=1}^{\mathrm{T}} \beta_j [w^{(j)}]^m$, it yields from (2.3) that

$$
\begin{aligned}
\langle P^{[m]}(A), B\rangle &= \left\langle \sum_{i=1}^{s} \alpha_i [Pu^{(i)}]^m, \sum_{j=1}^{\mathrm{T}} \beta_j [w^{(j)}]^m \right\rangle \\
&= \sum_{i=1}^{s}\sum_{j=1}^{\mathrm{T}} \alpha_i\beta_j \langle [Pu^{(i)}]^m, [w^{(j)}]^m\rangle \\
&= \sum_{i=1}^{s}\sum_{j=1}^{\mathrm{T}} \alpha_i\beta_j ((w^{(j)})^{\mathrm{T}} P u^{(i)})^m \\
&= \sum_{i=1}^{s}\sum_{j=1}^{\mathrm{T}} \alpha_i\beta_j ((u^{(i)})^{\mathrm{T}} P^{\mathrm{T}} w^{(j)})^m \\
&= \sum_{i=1}^{s}\sum_{j=1}^{\mathrm{T}} \alpha_i\beta_j \langle [P^{\mathrm{T}} w^{(j)}]^m, [u^{(i)}]^m\rangle \\
&= \left\langle \sum_{i=1}^{s} \alpha_i [u^{(i)}]^m, \sum_{j=1}^{\mathrm{T}} \beta_j [P^{\mathrm{T}} w^{(j)}]^m \right\rangle \\
&= \langle A, (P^{\mathrm{T}})^{[m]}(B)\rangle.
\end{aligned}
$$

This proves the assertion in (iii). The statement in (iv) follows from (ii) and (iii), and (v) and (vi) follow from (iii). □

We now characterize the range and null spaces of linear operator $P^{[m]}$, whose proofs are straightforward and we omit them here. They are basically determined by the range and null spaces of matrix P due to Theorem 2.2.

Proposition 2.4. *For any $P \in \mathbb{R}^{n\times n}$,*

(i) *the range space and the null space of $P^{[m]}$ are given by*

$$\mathcal{R}(P^{[m]}) = \left\{ \sum_{i=1}^{r} \alpha_i [u^{(i)}]^m : \alpha_i \in \mathbb{R}, u^{(i)} \in \mathcal{R}(P), i = 1, 2, \dots, r, r \in N \right\},$$
$$\mathcal{N}(P^{[m]}) = \left\{ \sum_{i=1}^{r} \alpha_i [u^{(i)}]^m : \alpha_i \in \mathbb{R}, u^{(i)} \in \mathcal{N}(P), i = 1, 2, \dots, r, r \in N \right\}.$$

(ii) *if P is invertible, then $\mathcal{R}(P^{[m]}) = \mathcal{T}(m,n)$ and $\mathcal{N}(P^{[m]}) = \{0\}$.*

We are also able to give explicit expressions of the range and null spaces of $P^{[m]}$ for any $P \in \mathbb{R}^{n\times n}$ in terms of the generalized inverse of a matrix, instead of the symmetric outer product decomposition form presented above. This resembles the projection theory of regular matrices.

Proposition 2.5. *For any $P \in \mathbb{R}^{n\times n}$, let $P^\dagger$ be the generalized inverse of P. Then,*

(i) $(P^{[m]})^\dagger = (P^\dagger)^{[m]}$;

(ii) $\mathcal{R}(P^{[m]}) = (PP^\dagger)^{[m]}(\mathcal{T}(m,n))$, $\mathcal{N}(P^{[m]}) = (I - (P^\dagger P)^{[m]})(\mathcal{T}(m,n))$;

(iii) *for any $A \in \mathcal{T}(m,n)$ and the tensor equation $P^{[m]}(X) = A$, $X = (P^\dagger)^{[m]}(A) + (I - (P^\dagger P)^{[m]})(B)$, $B \in \mathcal{T}(m,n)$ is a general solution to the equation if it is consistent, or a least-squares solution if it is inconsistent.*

Actually, the class of linear operators defined in Definition 2.1 and its linear combinations play a fundamental role in the symmetric tensor space which can act as a subtensor generator, some special projection operators, diagonalization operators and even the Hadamard product, as the following proposition illustrates.

Proposition 2.6. *For any $B \in \mathcal{T}(m,n)$,*

(i) *let Γ_k be any index subset of $\{1, 2, \dots, n\}$ with $|\Gamma_k| = k$ $(k = 1, \dots, n)$, $B_{\Gamma_k} \in \mathcal{T}(m,k)$ be the corresponding m-order k-dimensional sub-tensor of B, and $\overline{B_{\Gamma_k}}$ be the natural expansion of B_{Γ_k}. Then $(I_{\Gamma_k})^{[m]}(B) = \overline{B_{\Gamma_k}}$, where $I_{\Gamma_k} \in \mathbb{R}^{n\times n}$ is the matrix with the (i,i)-entry 1 if $i \in \Gamma_k$ and other entries 0;*

(ii) *let $w^{(1)}, w^{(2)}, \dots, w^{(h)} \in \mathbb{R}^n$ be unit vectors which are pairwise orthogonal, and $S_{\{w^{(1)},w^{(2)},\dots,w^{(h)}\}} := \{\sum_{i=1}^h k_i [w^{(i)}]^m : k_i \in \mathbb{R}, i = 1, 2, \dots, h\}$. Then the projection of B onto $S_{\{w^{(1)},w^{(2)},\dots,w^{(h)}\}}$ is $\sum_{i=1}^h (P^{(i)})^{[m]}(B)$, with $P^{(i)} := w^{(i)}(w^{(i)})^{\mathrm{T}}$;*

(iii) *let $\mathrm{diag}(B)$ be the diagonal tensor which remains the diagonal entries of B and others 0. Then $\mathrm{diag}(B) = \sum_{i=1}^n (\tilde{P}^{(i)})^{[m]}(B)$, with $\tilde{P}^{(i)} := e^{(i)}(e^{(i)})^{\mathrm{T}}$, where $e^{(i)} \in \mathbb{R}^n$ with 1 at the ith entry and 0 everywhere else;*

(iv) *let $A := \sum_{i=1}^s \alpha_i [u^{(i)}]^m$ and $A \circ B$ be the Hadamard product of A and B. Then $A \circ B = \sum_{i=1}^s \alpha_i (\bar{P}^{(i)})^{[m]}(B)$, with $\bar{P}^{(i)}$ the diagonal matrix generated by $u^{(i)}$.*

Proof. (i), (ii) and (iii) follow by definition. For any $B \in \mathcal{T}(m,n)$ with any of its symmetric outer product decomposition $B = \sum_{j=1}^t \beta_j [v^{(j)}]^m$, it follows from the decomposition invariance property that for any $i_1, \dots, i_m \in \{1, 2, \dots, n\}$,

$$\begin{aligned}
\left(\sum_{i=1}^s \alpha_i (\bar{P}^{(i)})^{[m]}(B) \right)_{i_1\cdots i_m} &= \left(\sum_{i=1}^s \alpha_i \sum_{j=1}^t \beta_j [\bar{P}^{(i)} v^{(j)}]^m \right)_{i_1\cdots i_m} \\
&= \sum_{i=1}^s \alpha_i \sum_{j=1}^t \beta_j (u^{(i)})_{i_1} (v^{(j)})_{i_1} \cdots (u^{(i)})_{i_m} (v^{(j)})_{i_m} \\
&= \left(\sum_{i=1}^s \alpha_i (u^{(i)})_{i_1} \cdots (u^{(i)})_{i_m} \right) \left(\sum_{j=1}^t \beta_j (v^{(j)})_{i_1} \cdots (v^{(j)})_{i_m} \right)
\end{aligned}$$

$$= A_{i_1\cdots i_m} B_{i_1\cdots i_m}.$$

Thus the assertion in (iv) follows and the whole proof is complete. □

3 Positive semidefiniteness of linear operators

A linear operator is said to be positive semidefinite if it is monotone and self-adjoint. It is known that the monotonicity of the gradient function leads to the convexity of the primitive real-valued function and the self-adjointness of a linear operator indicates some symmetry property with respect to the inner product. This section is devoted to the positive semidefiniteness of aforementioned linear operators in symmetric tensor spaces. We begin with a basic theorem on such issue as follows.

Theorem 3.1. *For any $P \in \mathbb{R}^{n\times n}$, $(P^{\mathrm{T}}P)^{[m]}$ is a positive semidefinite operator in $\mathcal{T}(m,n)$.*

Proof. For any $A \in \mathcal{T}(m,n)$ with any of its symmetric outer product decomposition $A = \sum_{i=1}^{s} \alpha_i [u^{(i)}]^m$, by the decomposition invariance property, we have

$$\begin{aligned}
\langle (P^{\mathrm{T}}P)^{[m]}(A), A\rangle &= \left\langle \sum_{i=1}^{s} \alpha_i [P^{\mathrm{T}}Pu^{(i)}]^m, \sum_{j=1}^{s} \alpha_j [u^{(j)}]^m \right\rangle \\
&= \sum_{i=1}^{s}\sum_{j=1}^{s} \alpha_i\alpha_j \langle P^{\mathrm{T}}Pu^{(i)}, u^{(j)}\rangle^m \\
&= \sum_{i=1}^{s}\sum_{j=1}^{s} \alpha_i\alpha_j \langle Pu^{(i)}, Pu^{(j)}\rangle^m \\
&= \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_s \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \langle Pu^{(1)}, Pu^{(1)}\rangle^m & \cdots & \langle Pu^{(1)}, Pu^{(s)}\rangle^m \\ \langle Pu^{(2)}, Pu^{(1)}\rangle^m & \cdots & \langle Pu^{(2)}, Pu^{(s)}\rangle^m \\ \vdots & \ddots & \vdots \\ \langle Pu^{(s)}, Pu^{(1)}\rangle^m & \cdots & \langle Pu^{(s)}, Pu^{(s)}\rangle^m \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_s \end{bmatrix} \\
&=: w^{\mathrm{T}} M w,
\end{aligned}$$

where $w := (\alpha_1, \alpha_2, \ldots, \alpha_s)^{\mathrm{T}} \in \mathbb{R}^s$ and

$$M := \begin{bmatrix} \langle Pu^{(1)}, Pu^{(1)}\rangle^m & \cdots & \langle Pu^{(1)}, Pu^{(s)}\rangle^m \\ \langle Pu^{(2)}, Pu^{(1)}\rangle^m & \cdots & \langle Pu^{(2)}, Pu^{(s)}\rangle^m \\ \vdots & \ddots & \vdots \\ \langle Pu^{(s)}, Pu^{(1)}\rangle^m & \cdots & \langle Pu^{(s)}, Pu^{(s)}\rangle^m \end{bmatrix} \in \mathbb{R}^{s\times s}.$$

Denote $U := [u^{(1)}, \ldots, u^{(s)}] \in \mathbb{R}^{n\times s}$. It is easy to verify that

$$M = \underbrace{((PU)^{\mathrm{T}}PU) \circ ((PU)^{\mathrm{T}}PU) \circ \cdots \circ ((PU)^{\mathrm{T}}PU)}_{m \text{ copies}},$$

where $\circ$ is the Hadamard product. Noting that $((PU)^{\mathrm{T}}PU) \in S(2,s)$, together with [11, Theorem 5.2.1], we have $M \in \mathcal{S}(2,s)$ for any $U \in \mathbb{R}^{n\times s}$, which further implies that $w^{\mathrm{T}}Mw \geqslant 0$ for any $w \in \mathbb{R}^s$. Thus $\langle (P^{\mathrm{T}}P)^{[m]}(A), A\rangle \geqslant 0$ for any $A \in \mathcal{T}(m,n)$ and hence $(P^{\mathrm{T}}P)^{[m]}$ is positive semidefinite. □

When $\mathcal{T}(m,n)$ is specified to the symmetric matrices space, Theorem 3.1 turns out to be the fact that for any $\Psi := P^{\mathrm{T}}P$, the Kronecker product of Ψ and itself, denoted by K_Ψ, remains symmetric positive semidefinite, as we can easily verify that the vectorization of $\Psi X \Psi$ is exactly $K_\Psi \mathrm{vec}(X)$, where $\mathrm{vec}(\cdot)$ is the vectorization operator. Theorem 3.1 further implies the convexity of the Frobenius norm function $f(A) := \frac{1}{2}\|P^{[m]}(A)\|_F^2$ in tensor A.

Proposition 3.2. *For any given $P \in \mathbb{R}^{n\times n}$, the gradient of $f(A) := \frac{1}{2}\|P^{[m]}(A)\|_F^2$ at A is $\nabla f(A) = (P^{\mathrm{T}}P)^{[m]}(A)$. Moreover, $f(A)$ is a convex function of A.*

Proof. By direct calculation, we have

$$f(A) = \frac{1}{2}\sum_{i_1\cdots i_m=1}^{n}\Bigg(\sum_{i'_1,\dots,i'_m=1}^{n} P_{i_1 i'_1}\cdots P_{i_m i'_m} A_{i'_1\cdots i'_m}\Bigg)^2,$$

which further implies that

$$\begin{aligned}
\frac{\partial f}{\partial A_{j_1\cdots j_m}} &= \sum_{i_1\cdots i_m=1}^{n}\Bigg(\sum_{i'_1,\dots,i'_m=1}^{n} P_{i_1 i'_1}\cdots P_{i_m i'_m} A_{i'_1\cdots i'_m}\Bigg)P_{i_1 j_1}\cdots P_{i_m j_m}\\
&= \sum_{i_1\cdots i_m=1}^{n} (P^{[m]}(A))_{i_1\cdots i_m} P_{i_1 j_1}\cdots P_{i_m j_m}\\
&= ((P^{\mathrm{T}})^{[m]}(P^{[m]}(A)))_{j_1\cdots j_m}\\
&= ((P^{\mathrm{T}}P)^{[m]}(A))_{j_1\cdots j_m}.
\end{aligned}$$

Henceforth, $\nabla f(A) = (P^{\mathrm{T}}P)^{[m]}(A)$. Applying Theorem 3.1, together with the first order property of convex function in convex analysis, we get the convexity of $f(A)$ immediately. □

Qi and Ye [18] proposed a linear operator $\mathcal{L}$ which builds up a one-to-one correspondence from $\mathcal{T}(m,3)$ to $\mathbb{R}^\kappa$ with $\kappa = \frac{(m+2)(m+1)}{2}$. They further managed to achieve the equivalence between positive semidefinite linear operators on $\mathcal{T}(m,3)$ and symmetric positive semidefinite matrices in $\mathbb{R}^{\kappa\times\kappa}$. Similarly, we restrict our attention also to 3-dimensional symmetric tensors, which are called space tensors; see [18], and consider the positive semidefiniteness of some special linear operators via the corresponding matrix representations in the remaining of this section.

Lemma 3.3. *Let $\mathcal{Q} := \sum_{i=1}^h \alpha_i (u^{(i)}(u^{(i)})^{\mathrm{T}})^{[m]}$ with $u^{(i)} \in \mathbb{R}^3$. Then the corresponding matrix representation of $\mathcal{Q}$ is*

$$Q := \sum_{i=1}^{h} \alpha_i \mathcal{L}([u^{(i)}]^m)\mathcal{L}([u^{(i)}]^m)^{\mathrm{T}}.$$

Proof. Note that for any $A \in \mathcal{T}(m,3)$, we have

$$\begin{aligned}
\mathcal{L}(\mathcal{Q}(A)) &= \mathcal{L}\Bigg(\sum_{i=1}^{h}\alpha_i (u^{(i)}(u^{(i)})^{\mathrm{T}})^{[m]}(A)\Bigg)\\
&= \mathcal{L}\Bigg(\sum_{i=1}^{h}\alpha_i \langle [u^{(i)}]^m, A\rangle [u^{(i)}]^m\Bigg)\\
&= \sum_{i=1}^{h}\alpha_i \langle [u^{(i)}]^m, A\rangle \mathcal{L}([u^{(i)}]^m)\\
&= \sum_{i=1}^{h}\alpha_i \langle \mathcal{L}([u^{(i)}]^m), \mathcal{L}(A)\rangle \mathcal{L}([u^{(i)}]^m)\\
&= \sum_{i=1}^{h}\alpha_i \mathcal{L}([u^{(i)}]^m)\mathcal{L}([u^{(i)}]^m)^{\mathrm{T}}\mathcal{L}(A).
\end{aligned}$$

In view of the identity $\mathcal{L}(\mathcal{Q}(A)) = Q(\mathcal{L}(A))$, the desired result follows. □

Based on the lemma, we consider the rank of the representation matrix.

Proposition 3.4. *Let $\mathcal{Q} := \sum_{i=1}^h \alpha_i (u^{(i)}(u^{(i)})^{\mathrm{T}})^{[m]}$ with $u^{(1)}, u^{(2)}, \dots, u^{(h)} \in \mathbb{R}^3$ pairwise linearly independent, and $m \geqslant h-1$. Then the rank of the corresponding matrix representation Q is h if all $\alpha_i \neq 0$, and $\mathcal{Q}$ is positive semidefinite iff $\alpha_i \geqslant 0$ for all $i = 1, 2, \dots, h$.*

Proof. By the fact that $u^{(1)}, u^{(2)}, \dots, u^{(h)} \in \mathbb{R}^3$ are pair-wise linearly independent, we have $[u^{(1)}]^m$, $[u^{(2)}]^m, \dots, [u^{(h)}]^m$ are linearly independent from [6, Corollary 4.4]. Note that $\mathcal{L}$ is the linear isomorphic operator between $\mathcal{T}(m,3)$ and $\mathbb{R}^\kappa$ with $\kappa := \frac{1}{2}(m+1)(m+2)$. Thus the linear independency of $\mathcal{L}([u^{(1)}]^m)$, $\mathcal{L}([u^{(2)}]^m), \dots, \mathcal{L}([u^{(h)}]^m)$ follows. This further implies that $\mathcal{L}([u^{(1)}]^m)\mathcal{L}([u^{(1)}]^m)^{\mathrm{T}}$, $\mathcal{L}([u^{(2)}]^m)\mathcal{L}([u^{(2)}]^m)^{\mathrm{T}}, \dots, \mathcal{L}([u^{(h)}]^m)\mathcal{L}([u^{(h)}]^m)^{\mathrm{T}}$ are linearly independent. If $\alpha_i \neq 0$, then Q has rank h. The remaining part can be verified directly by definition. □

Note that if $h \leqslant m+1$, then $h \leqslant \kappa = \frac{1}{2}(m+1)(m+2)$ and hence the matrix representation of $\mathcal{Q}$ defined in the above proposition is always rank-deficient. Next, we give an explicit form of a linear operator which enjoys the positive definiteness like a regular matrix.

Proposition 3.5. *Let $v^{(1)}$, $v^{(2)}, \dots, v^{(\kappa)} \in \mathbb{R}^\kappa$ be any linearly independent unit vectors, $\bar{S}_{v^{(i)}} := \{k\mathcal{L}^{-1}(v^{(i)}) : k \in \mathbb{R}\}$ $(i = 1, 2, \dots, \kappa)$ and $\mathcal{Q} := \sum_{i=1}^t q_i \mathcal{P}_{\bar{S}_{v^{(i)}}}$ with $q_i \in \mathbb{R}$, where $\mathcal{P}_{\bar{S}_{v^{(i)}}}$ is the projection operator on $\bar{S}_{v^{(i)}}$. Here $\mathcal{L}^{-1}$ is the inverse operator of $\mathcal{L}$. Then the corresponding matrix representation is $Q := \sum_{i=1}^t q_i v^{(i)}(v^{(i)})^{\mathrm{T}}$. Moreover, $\mathcal{Q}$ is symmetric and*

(i) *$\mathcal{Q}$ is positive semidefinite iff all $q_i \geqslant 0$;*

(ii) *$\mathcal{Q}$ is positive definite iff all $q_i > 0$.*

Proof. We can verify by definition that

$$\mathcal{P}_{\bar{S}_{v^{(i)}}}(A) = \langle \mathcal{L}^{-1}(v^{(i)}), A\rangle \mathcal{L}^{-1}(v^{(i)}), \quad \forall\, A \in \mathcal{T}(m,3).$$

Therefore,

$$\begin{aligned}\mathcal{L}(\mathcal{Q}(A)) &= \mathcal{L}\bigg(\sum_{i=1}^t q_i \langle \mathcal{L}^{-1}(v^{(i)}), A\rangle \mathcal{L}^{-1}(v^{(i)})\bigg)\\ &= \sum_{i=1}^t q_i \langle \mathcal{L}^{-1}(v^{(i)}), A\rangle v^{(i)}\\ &= \sum_{i=1}^t q_i \langle v^{(i)}, \mathcal{L}(A)\rangle v^{(i)},\end{aligned}$$

which implies that $Q(\mathcal{L}(A)) = \sum_{i=1}^t q_i v^{(i)}(v^{(i)})^{\mathrm{T}}\mathcal{L}(A)$. Thus, the first part of the assertion is established. Then the "moreover" part follows directly by definition. □

Indeed, the projection $\mathcal{P}_{\bar{S}_{v^{(i)}}}(A)$ mentioned in the above proposition is exactly a kind of contraction product (see [7]) of the $2m$-order n-dimensional tensor $\mathcal{L}^{-1}(v^{(i)}) \otimes \mathcal{L}^{-1}(v^{(i)})$ and A, which can be specified as

$$\mathcal{P}_{\bar{S}_{v^{(i)}}}(A) = (\mathcal{L}^{-1}(v^{(i)}) \otimes \mathcal{L}^{-1}(v^{(i)})) \bullet_1 \bullet_2 \cdots \bullet_m A.$$

Here $\mathcal{L}^{-1}(v^{(i)}) \otimes \mathcal{L}^{-1}(v^{(i)})$ is the outer product of $\mathcal{L}^{-1}(v^{(i)})$ and itself, which is defined as

$$(\mathcal{L}^{-1}(v^{(i)}) \otimes \mathcal{L}^{-1}(v^{(i)}))_{i_1\cdots i_m j_1 \cdots j_m} = (\mathcal{L}^{-1}(v^{(i)}))_{i_1\cdots i_m}(\mathcal{L}^{-1}(v^{(i)}))_{j_1\cdots j_m}.$$

Obviously, such $2m$-order tensor loses the symmetry for any $m \geqslant 2$.

4 Symmetric positive semidefinite tensor cone

Analogous with the positive semidefinite cone in the symmetric matrix space, we now turn our attention to the positive semidefinite tensor cone in the symmetric tensor space. Positive definiteness is an intrinsic property. For many real applications, a tensor must be positive semidefinite to be physically meaningful. In this section, we characterize the self-duality of the SDT cone and some face structures of its dual cone by employing the properties of aforementioned linear operators. The high-order SDT cone, first defined by Qi and Ye [18] for space tensors, will be extended to the space $\mathcal{T}(m,n)$ with any $n \geqslant 2$ in this paper. First, we define what the SDT cone is.

Definition 4.1. For any $A \in \mathcal{T}(m,n)$, we say A is positive semidefinite if

$$A[x]^m \geqslant 0, \quad \forall\, x \in \mathbb{R}^n.$$

The set of all positive semidefinite tensors in $\mathcal{T}(m,n)$, denoted by $\mathcal{S}(m,n)$, is called symmetric positive semidefinite tensor cone.

Similar to [18, Theorem 3], we can also derive the dual cone of $\mathcal{S}(m,n)$ as follows.

Proposition 4.2. *The dual cone of $\mathcal{S}(m,n)$, termed as $\mathcal{V}(m,n)$, has the form*:

$$\mathcal{V}(m,n) = \left\{ \sum_{i=1}^{l} [x^{(i)}]^m : x^{(i)} \in \mathbb{R}^n, i = 1, 2, \dots, l \right\},$$

where l is the dimension of the space $\mathcal{T}(m,n)$.

Evidently, $\mathcal{V}(m,n) \subseteq \mathcal{S}(m,n)$. In particular, when $m = 2$, both $\mathcal{S}(m,n)$ and $\mathcal{V}(m,n)$ coincide as the well-known symmetric positive semidefinite matrix cone which plays a fundamental role in semidefinite programming [2, 21]. Thus, $\mathcal{S}(2,n) = \mathcal{V}(2,n)$ and so that it is self dual. The natural question would be: Can the self-duality be extended to $m \geqslant 4$? We show next that the self-duality is not true for $m \geqslant 4$. Before we proceed, we characterize some faces of $\mathcal{V}(m,n)$. It is evident that the set $S_w^+ := \{k[w]^m : k \in \mathbb{R}, k \geqslant 0\}$ is an extreme ray with any given $w \in \mathbb{R}^n$ and hence a face of $\mathcal{V}(m,n)$ and $S(m,n)$. Next we give another class of faces of $\mathcal{V}(m,n)$.

Theorem 4.3. *For any $k \in \{1, \dots, n\}$ and any index subset $\Gamma_k \subseteq \{1, \dots, n\}$ with $|\Gamma_k| = k$, let I_{Γ_k} be defined as in Proposition* 2.6(i). *Then we have*

$$(I_{\Gamma_k})^{[m]}(\mathcal{V}(m,n)) \trianglelefteq \mathcal{V}(m,n).$$

Proof. Without loss of generality, we may assume that $\Gamma_k = \{1, 2, \dots, k\}$ with any given $1 \leqslant k \leqslant n$. For any $A, B \in \mathcal{V}(m,n)$, with $A = \sum_{j=1}^{r_1} [u^{(j)}]^m$ and $B = \sum_{l=1}^{r_2} [w^{(l)}]^m$. If $A + B \in (I_{\Gamma_k})^{[m]}(\mathcal{V}(m,n))$, then for any $x \in \mathbb{R}^n$ with $x_i = 0$ for any $i \in \{1, \dots, k\}$, $(A+B)[x]^m = 0$. Combining with the fact A, $B \in \mathcal{V}(m,n) \subseteq \mathcal{S}(m,n)$, we achieve that $A[x]^m = B[x]^m = 0$. By setting $x^{(i)} \in \mathbb{R}^n$ with $x^{(i)}_{i+k} \neq 0$ and other components 0 for any $i = 1, \dots, n-k$, we can then obtain that

$$\sum_{j=1}^{r_1} (u^{(j)})^m_{i+k} (x^{(i)}_{i+k})^m = \sum_{l=1}^{r_2} (w^{(l)})^m_{i+k} (x^{(i)}_{i+k})^m = 0, \quad \forall\, i = 1, \dots, n-k,$$

which further implies that for any $i = 1, \dots, n-k$,

$$(u^{(j)})_{i+k} = 0, \quad \forall\, j = 1, \dots, r_1,$$
$$(w^{(l)})_{i+k} = 0, \quad \forall\, l = 1, \dots, r_2.$$

Henceforth, $A, B \in (I_{\Gamma_k})^{[m]}(\mathcal{V}(m,n))$. By definition, we establish that $(I_{\Gamma_k})^{[m]}(\mathcal{V}(m,n))$ is a face of $\mathcal{V}(m,n)$. □

Utilizing the similar technique in the above proof, we can also get the following result which is analogous with the one of symmetric positive semidefinite matrices.

Proposition 4.4. *For any $A \in \mathcal{V}(m,n)$, if there exists an index $i \in \{1, 2, \dots, n\}$ such that $A_{ii\cdots i} = 0$, then $A \in (I_{\Gamma_{n-1}})^{[m]}(\mathcal{V}(m,n))$ with $\Gamma_{n-1} = \{1, \dots, i-1, i+1, \dots, n\}$. Particularly, if all diagonal elements of A are* 0, *then $A = 0$.*

Proof. Since $A \in \mathcal{V}(m,n)$, we can find some integer s and $u_1, \dots, u_s \in \mathbb{R}^n$ such that $A = \sum_{j=1}^{s} [u^{(j)}]^m$. Note that

$$A_{ii\cdots i} = (u^{(1)})_i^m + \cdots + (u^{(s)})_i^m = 0.$$

Together with the fact that m is even and all $u^{(i)} \in \mathbb{R}^n$, it follows readily that

$$(u^{(1)})_i = (u^{(2)})_i = \cdots = (u^{(s)})_i = 0,$$

which implies that $A_{i_1\cdots i_m} = 0$ if $i \in \{i_1, \ldots, i_m\}$. Henceforth, we have $A \in (I_{\Gamma_{n-1}})^{[m]}(\mathcal{V}(m,n))$ with $\Gamma_{n-1} = \{1, \ldots, i-1, i+1, \ldots, n\}$ from Proposition 2.6(i). Similarly, we can verify the second part of the proposition. This completes the proof. □

The above property may fail to hold for $A \in \mathcal{S}(m,n)$, as the following example shows.

Example 4.5. Let $A \in \mathcal{T}(4,3)$ with $A_{1122} = A_{1212} = A_{1221} = A_{2112} = A_{2121} = A_{2211} = 1$ and other entries 0. Then for any $x = (x_1, x_2, x_3)^{\mathrm{T}} \in \mathbb{R}^3$, we have $A[x]^4 = 6x_1^2x_2^2 \geqslant 0$. Thus $A \in \mathcal{S}(4,3)$.

The above example also shows that $\mathcal{S}(4,3) \neq \mathcal{V}(4,3)$, since the aforementioned $A \notin \mathcal{V}(4,3)$ by Proposition 4.4. This further comes to the self-duality result as follows.

Theorem 4.6. *For any $n \geqslant 2$, we have $\mathcal{V}(m,n) \subseteq \mathcal{S}(m,n)$, and the equality holds if and only if $m = 2$. Furthermore, $\mathrm{int}(\mathcal{V}(m,n)) \subseteq \mathrm{int}(\mathcal{S}(m,n))$, and the equality holds if and only if $m = 2$.*

Proof. The inclusion is evident. Thus it remains to show the equivalence between the equality and $m = 2$. The sufficiency is valid by the self-duality of positive semidefinite matrices cone. To get the necessity, it suffices to construct a symmetric tensor A which lies in $\mathcal{S}(m,n) \setminus \mathcal{V}(m,n)$ for any $m \geqslant 4$. Let $A \in \mathcal{T}(m,n)$ with $A_{1122\cdots 2} = A_{1212\cdots 2} = \cdots = A_{2\cdots 211} = 1$ and other entries 0. It is easy to verify that

$$A[x]^m = \alpha x_1^2 x_2^{m-2} \geqslant 0, \quad \forall\, x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n,$$

where $\alpha = \frac{m(m-1)}{2}$. Thus $A \in \mathcal{S}(m,n)$. However, such a tensor A has all its diagonal elements 0 and henceforth $A \notin \mathcal{V}(m,n)$ by Proposition 4.4. Together with the closedness of $\mathcal{V}(m,n)$ and $\mathcal{S}(m,n)$, we can obtain the relation of their interiors. This completes the whole proof. □

5 Relationship of positive semidefinite tensor cones

Followed by Theorem 4.6, we move on to considering some special subsets of $\mathrm{int}(\mathcal{S}(m,n)) \setminus \mathrm{int}(\mathcal{V}(m,n))$, i.e., subsets between $\mathcal{S}(m,n)$ and $\mathcal{V}(m,n)$ for any $m \geqslant 4$ and further exhibit the relationship to other cones in tensor spaces. We will start with a key lemma as a theoretical preparation.

Lemma 5.1. *We can express the interiors of $\mathcal{S}(m,n)$ and $\mathcal{V}(m,n)$, respectively as*

$$\mathrm{int}(\mathcal{S}(m,n)) = \{A \in \mathcal{S}(m,n) : \langle A, B\rangle > 0, \forall B \in \mathcal{V}(m,n), B \neq 0\},$$
$$\mathrm{int}(\mathcal{V}(m,n)) = \{B \in \mathcal{V}(m,n) : \langle A, B\rangle > 0, \forall A \in \mathcal{S}(m,n), A \neq 0\}.$$

Proposition 5.2. *For any even $m \geqslant 4$, letting $\bar{D}(m,n)$ be the set of all diagonal tensors in $\mathcal{T}(m,n)$ with all diagonal elements positive, we have $\bar{D}(m,n) \subseteq \mathrm{int}(\mathcal{S}(m,n)) \setminus \mathrm{int}(\mathcal{V}(m,n))$.*

Proof. For any diagonal tensor B with the corresponding diagonal elements $\alpha_1, \ldots, \alpha_n > 0$, we have

$$B[x]^m = \sum_{i=1}^{n} \alpha_i x_i^m > 0, \quad \forall\, x \in \mathbb{R}^n, \quad x \neq 0.$$

Thus, $B \in \mathrm{int}(\mathcal{S}(m,n))$. However, by setting a nonzero tensor $A \in \mathcal{S}(m,n)$ as defined in the proof of Theorem 4.6, we can get $\langle A, B\rangle = 0$. In view of Lemma 5.1, it follows that $B \notin \mathrm{int}(\mathcal{V}(m,n))$. Henceforth, the desired result is established. □

The aforementioned diagonal tensor with positive diagonal entries is exactly a special case of nonnegative tensors which have numerous applications such as high-order Markov chains [16], spectral hyper-graph theory [4], multi-linear pagerank in the internet [14]. This inspires us to exploit more connections between nonnegative tensors and positive semidefinite tensors. As a start, a generalization of diagonal tensors is introduced as follows.

Definition 5.3. For any $A \in \mathcal{T}(m,n)$, the entries $A_{i_1\cdots i_m}$ whose subscript $i_1\cdots i_m$ can be written as $j_1j_1\cdots j_{m/2}j_{m/2}$ by permutation, are called the partially diagonal entries of A, others are called the off partially diagonal entries. A tensor $A \in \mathcal{T}(m,n)$ is said to be partially diagonal if all its off partially diagonal entries are 0.

Actually, the partially diagonal tensors in $\mathcal{T}(m,n)$ can be easily verified to be quasi-diagonal tensors which were defined by Chang et al. [5]. Denote the set of all partially diagonal tensors in $\mathcal{T}(m,n)$ by $D(m,n)$, and the set of all tensors in $D(m,n)$ with nonnegative (or respectively, positive) partially diagonal entries by $D_+(m,n)$ (or respectively, $D_{++}(m,n)$). It is evident that all diagonal tensors are partially diagonal and when $m=2$, $D(m,n)$ is exactly the set of all diagonal matrices in $\mathbb{R}^{n\times n}$.

Proposition 5.4. *$D_+(m,n)$ is a pointed, convex, closed cone in $\mathcal{T}(m,n)$ and*

(i) *$D_+(m,n)$ has no interior in $\mathcal{T}(m,n)$;*

(ii) *$D_+(m,n)$ has the interior $D_{++}(m,n)$ in $D(m,n)$.*

The relation among $\mathcal{S}(m,n)$, $\mathcal{V}(m,n)$ and $D_+(m,n)$ are as follows.

Theorem 5.5. *For any $n\geqslant 2$, we have*

(i) *$D_+(m,n)\subset\mathcal{S}(m,n)$, and $D_{++}(m,n)\subset\operatorname{int}(\mathcal{S}(m,n))$;*

(ii) *Every tensor in $\mathcal{V}(m,n)$ has nonnegative partially diagonal entries.*

Before providing the proof of the above theorem, we will first introduce another type of sub-tensors in $\mathcal{T}(m,n)$. For any $A=\sum_{i=1}^s\alpha_i[u^{(i)}]^m$, we say

$$A^{(kl)}_{(m-2)}=\sum_{i=1}^s\alpha_i[u^{(i)}]^{(m-2)}(u^{(i)})_k(u^{(i)})_l$$

is an $(m-2)$-order n-dimensional subtensor of A. There are $\frac{n(n+1)}{2}$ different $(m-2)$-order n-dimensional subtensors of A. By definition, we can immediately get that for any $A\in\mathcal{T}(m,n)$,

$$(A^{(kl)}_{(m-2)})_{i_1\cdots i_{m-2}}=A_{i_1\cdots i_{m-2}kl},\quad\forall\, i_1,\ldots,i_{m-2}\in\{1,2,\ldots,n\},$$

and hence

$$A[x]^m=\sum_{i_1,\ldots,i_{m-2}=1}^n\sum_{k,l=1}^n(A^{(kl)}_{(m-2)})_{i_1\cdots i_{m-2}}x_kx_lx_{i_1}\cdots x_{i_{m-2}},\quad\forall\, x\in\mathbb{R}^n.$$

Proposition 5.6. *For any $A\in\mathcal{V}(m,n)$, i.e., $A=\sum_{i=1}^s[u^{(i)}]^m$ with $s\leqslant\frac{(m+n-1)!}{m!(n-1)!}$, we have $A^{(kk)}_{(m-2)}\in\mathcal{V}(m-2,n)$ for any $k=1,2,\ldots,n$.*

Proof. By definition we have

$$A^{(kk)}_{(m-2)}=\sum_{i=1}^s[u^{(i)}]^{(m-2)}(u^{(i)})_k^2=\sum_{i=1}^s[(u^{(i)})_k^{\frac{2}{m-2}}u^{(i)}]^{(m-2)}.$$

This completes the proof. □

Proof of Theorem 5.5. From the above proposition, we can get $\frac{(n+\frac{m}{2}-1)!}{\frac{m}{2}!n!}$ different $n\times n$ symmetric positive semidefinite matrices for any $A\in\mathcal{V}(m,n)$ and each diagonal element of such matrices is exactly a partially diagonal entry of A. Noting that all diagonal elements of symmetric positive semidefinite matrices are nonnegative, it comes to the assertion of (ii). For any $A\in D_+(m,n)$, we have

$$A[x]^m=\sum_{j_1,\ldots,j_{m/2}}^n\alpha_{j_1,\ldots,j_{m/2}}A_{j_1j_1\cdots j_{m/2}j_{m/2}}x_{j_1}^2\cdots x_{j_{m/2}}^2\geqslant 0,\quad\forall\, x\in\mathbb{R}^n,$$

where $\alpha_{j_1,\ldots,j_{m/2}}>0$ is the number of different permutations of $j_1j_1\cdots j_{m/2}j_{m/2}$. This implies that $D_+(m,n)\subseteq\mathcal{S}(m,n)$. Let A be the tensor in $\mathcal{T}(m,n)$ with

$$A_{11\cdots 1}=1,\quad A_{22\cdots 2}=\frac{\alpha^2}{4},\quad A_{\underbrace{1\cdots 1}_{\frac{m}{2}\text{ copies}}\underbrace{2\cdots 2}_{\frac{m}{2}\text{ copies}}}=\cdots=A_{2\cdots 21\cdots 1}=-1,$$

where α is permutation number of the $1\cdots 12\cdots 2$. By the construction of A, we have $A\notin D_+(m,n)$, but

$$A[x]^m=x_1^m+\frac{\alpha^2}{4}x_2^m-\alpha x_1^{m/2}x_2^{m/2}\geqslant 0,\quad\forall\, x\in\mathbb{R}^n,$$

which implies that $A \in \mathcal{S}(m,n)$. Thus $D_+(m,n) \subset \mathcal{S}(m,n)$. For any $B \in D_{++}(m,n)$, it is easy to verify that

$$A[x]^m \geqslant \sum_{i=1}^n A_{ii\cdots i} x_i^m > 0, \quad \forall\, x \in \mathbb{R}^n, \quad x \neq 0.$$

This deduces that $D_{++}(m,n) \subset \mathrm{int}(\mathcal{S}(m,n))$. □

Applying Proposition 2.18 and Theorem 5.5(i), we can provide a class of tensors which belong to $\mathcal{S}(m,n)$ but not to $\mathcal{V}(m,n)$ for any $m \geqslant 4$ and $n \geqslant 2$.

Corollary 5.7. *For any $m \geqslant 4$ and $n \geqslant 2$, we have*

$$D_0(m,n) := \{A \in D_+(m,n) : A \neq 0, A_{ii\cdots i} = 0, \forall\, i = 1,2,\ldots,n\} \subseteq \mathcal{S}(m,n) \setminus \mathcal{V}(m,n).$$

From Theorem 5.5(ii), we know that $\mathcal{V}(m,n) \cap D_+(m,n) \neq \emptyset$, and from Example 4.5, we have $D_+(m,n) \setminus \mathcal{V}(m,n) \neq \emptyset$. It is easy to verify the non-emptiness of $\mathcal{V}(m,n) \setminus D_+(m,n)$. For example, let $A = [x]^4$ with $x = (1,-1)^{\mathrm{T}} \in \mathbb{R}^2$. By definition we know that $A \in \mathcal{V}(4,2)$ and by direct calculation we have all its partially diagonal entries 1 and off partially diagonal entries -1, which implies that $A \notin D_+(4,2)$. Henceforth, the cones $\mathcal{V}(m,n)$ and $D_+(m,n)$ properly intersect with each other.

It is known that any tensor $A \in \mathcal{T}(m,n)$ has a one-to-one m-degree homogeneous polynomial counterpart $f_A(x)$ as stated in (1.1). This correspondence further leads to equivalence that $f_A(x)$ is positive semidefinite iff $A \in \mathcal{S}(m,n)$. Inspired by such observation, we can similarly get a closed convex cone in $\mathcal{T}(m,n)$, termed as $\mathrm{SOS}(m,n)$, which is corresponding to all the m-order homogeneous polynomials which can be written as sum-of-squares. We employ the name in the literature of polynomial theory and call such cone the SOS *cone*.

Proposition 5.8. *For any $m \geqslant 4$ and $n \geqslant 2$, we have*

(i) $\mathcal{V}(m,n) \subset SOS(m,n) \subseteq \mathcal{S}(m,n)$;

(ii) $D_+(m,n) \subset SOS(m,n)$;

(iii) *The dual cone* $SOS(m,n)^* = \{B \in \mathcal{T}(m,n) : \langle B, C \otimes C\rangle \geqslant 0, \forall\, C \in \mathcal{T}(\frac{m}{2}, n)\}$ *and all partially diagonal entries of any tensor in* $SOS(m,n)^*$ *are nonnegative.*

The polynomial defined in (1.1) is a continuously differentiable function in $x \in \mathbb{R}^n$. By convex analysis, we know that $f_A(x)$ is convex in x if and only if $\nabla^2 f_A(x) \in \mathcal{S}(2,n)$ for any $x \in \mathbb{R}^n$. A new notion called SOS-convexity has recently been proposed as a tractable sufficient condition for convexity of polynomials based on sum of squares decomposition [1]. The definition is as follows.

Definition 5.9 (SOS-convexity [1])**.** A multivariate polynomial $f(x) \in \mathbb{R}[x_1,\ldots,x_n]$ is said to be SOS-convex if its Hessian matrix $H(x)$ can be factored as $H(x) = V(x)^{\mathrm{T}} V(x)$ for any $x \in \mathbb{R}^n$ with some $V(x) \in \mathbb{R}^{n\times n}$.

Proposition 5.10. *Let $f_A(x)$ be defined as in* (1.1). *We have*

(i) *if $f_A(x)$ is convex, then $A \in \mathcal{S}(m,n)$;*

(ii) *if $f_A(x)$ is SOS-convex, then $A \in SOS(m,n)$;*

(iii) *if $A \in \mathcal{V}(m,n)$, then $f_A(x)$ is SOS-convex and hence convex.*

Proof. By direct calculation, we have

$$\nabla^2 f_A(x) = m(m-1)A[x]^{m-2}, \quad \forall\, x \in \mathbb{R}^n. \tag{5.1}$$

By the second order property of convex function, it follows that $f_A(x)$ is convex iff

$$\nabla^2 f_A(x) \in \mathcal{S}(2,n), \quad \forall\, x \in \mathbb{R}^n.$$

Combining with (5.1), we obtain that for any $x \in \mathbb{R}^n$, $A[x]^{m-2} \in \mathcal{S}(2,n)$, which further implies that

$$A[x]^m = \langle x, A[x]^{m-2} x\rangle \geqslant 0, \quad \forall\, x \in \mathbb{R}^n.$$

This proves the assertion in (i). By Definition 5.9, together with (5.1), we know that $f_A(x)$ is SOS-convex iff $A[x]^{m-2} = V(x)^{\mathrm{T}} V(x)$ for some $V(x) \in \mathbb{R}[x_1,\ldots,x_n]^{k\times n}$ with some integer k, which further deduces that $A[x]^m = \langle x, V(x)^{\mathrm{T}}V(x)x\rangle = \sum_{i=1}^{k}(V(x)x)_i^2$. This proves the assertion in (ii). For any $A \in \mathcal{V}(m,n)$ with $A = \sum_{i=1}^{s}[u^{(i)}]^m$, we have

$$A[x]^{m-2} = \sum_{i=1}^{s}(x^{\mathrm{T}}u^{(i)})^{m-2}u^{(i)}(u^{(i)})^{\mathrm{T}} =: M(x)^{\mathrm{T}}M(x),$$

where

$$M(x) := [(x^{\mathrm{T}}u^{(1)})^{\frac{m-2}{2}}, \dots, (x^{\mathrm{T}}u^{(s)})^{\frac{m-2}{2}}]^{\mathrm{T}}.$$

This implies that $f_A(x)$ is SOS-convex by definition. □

Proposition 5.10(ii) is an analog to the case of general polynomials as described by Helton and Nie [10, Lemma 8]. Note that for $m = 2$, the convexity and SOS-convexity coincide with each other for $f_A(x) = x^{\mathrm{T}}Ax$ and such a property holds iff A is positive semidefinite. However, for higher order cases, it is not the case, as the following statements show:

(i) if $A \in \mathcal{S}(m,n)$, the function $f_A(x)$ may not be convex;

(ii) if $A \in \mathrm{SOS}(m,n)$, the function $f_A(x)$ may not be SOS-convex;

(iii) if $f_A(x)$ is convex, the corresponding A may not lie in $\mathcal{V}(m,n)$.

Two counterexamples are proposed here to illustrate the above remark, where the second one is employed from [1].

Example 5.11. (i) Let $A \in \mathcal{T}(4,n)$ with $A_{1122} = A_{1212} = A_{1221} = A_{2112} = A_{2121} = A_{2211} = 1$ and other entries 0. Then for any $x = (x_1,\dots,x_n)^{\mathrm{T}} \in \mathbb{R}^3$, we have $A[x]^4 = 6x_1^2x_2^2 \geqslant 0$. Thus $A \in \mathrm{SOS}(m,n) \subseteq \mathcal{S}(4,n)$. However,

$$A[x]^2 = \begin{bmatrix} x_2^2 & 2x_1x_2 & 0 & \cdots & 0 \\ 2x_1x_2 & x_1^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

For any $x \in \mathbb{R}^n$, it is obvious that $A[x]^2$ may not lie in $\mathcal{S}(2,n)$ which implies that $f_A(x)$ is not convex and hence not SOS-convex.

(ii) Let $f_B(x) = 32x_1^8 + 118x_1^6x_2^2 + 40x_1^6x_3^2 + 25x_1^4x_2^4 - 43x_1^4x_2^2x_3^2 - 35x_1^4x_3^4 + 3x_1^2x_2^4x_3^2 - 16x_1^2x_2^2x_3^4 + 24x_1^2x_3^6 + 16x_2^8 + 44x_2^6x_3^2 + 70x_2^4x_3^4 + 60x_2^2x_3^6 + 30x_3^8$. It is shown in [1] that such 8-degree homogeneous polynomial $f_B(x)$ is convex. Note that the corresponding $B \in \mathcal{T}(8,3)$ has negative partially diagonal entries. Applying Theorem 5.5(ii), we have $B \notin \mathcal{V}(8,3)$.

To get more connections with the nonnegative tensor cone, we state a basic theorem for the linear images of the SDT cone and its dual cone as follows.

Theorem 5.12. *For any $P \in \mathbb{R}^{n\times n}$, we have*

(i) *$P^{[m]}(\mathcal{S}(m,n)) \subseteq \mathcal{S}(m,n)$ and the equality holds iff P is invertible*;

(ii) *$P^{[m]}(\mathcal{V}(m,n)) \subseteq \mathcal{V}(m,n)$ and the equality holds iff P is invertible.*

Proof. For any $x \in \mathbb{R}^n$, and any $A \in \mathcal{S}(m,n)$ with any of its symmetric outer product decomposition $A = \sum_{i=1}^{s}\alpha_i[u^{(i)}]^m$, it follows that

$$\begin{aligned} P^{[m]}(A)[x]^m &= \sum_{i=1}^{s}\alpha_i\langle[Pu^{(i)}]^m, [x]^m\rangle \\ &= \sum_{i=1}^{s}\alpha_i\langle Pu^{(i)}, x\rangle^m \\ &= \sum_{i=1}^{s}\alpha_i\langle u^{(i)}, P^{\mathrm{T}}x\rangle^m \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{s} \alpha_i \langle [u^{(i)}]^m, [P^{\mathrm{T}}x]^m \rangle \\
&= A[P^{\mathrm{T}}x]^m \\
&\geqslant 0,
\end{aligned}
$$

which reveals that $P^{[m]}(A) \in \mathcal{S}(m,n)$. Thus,

$$P^{[m]}(\mathcal{S}(m,n)) \subseteq \mathcal{S}(m,n).$$

Combining with Proposition 2.3(ii), we can obtain under the invertibility of P that

$$\mathcal{S}(m,n) = P^{[m]}(P^{-1})^{[m]}(\mathcal{S}(m,n)) \subseteq P^{[m]}(\mathcal{S}(m,n)) \subseteq \mathcal{S}(m,n),$$

which is equivalent to $P^{[m]}(\mathcal{S}(m,n)) = \mathcal{S}(m,n)$. If P is not invertible, then there exists some $x \in \mathbb{R}^n \setminus \{0\}$ such that $Px = 0$. It is easy to verify that

$$[x]^m \in \mathcal{S}(m,n) \setminus P^{[m]}(\mathcal{S}(m,n)).$$

Henceforth, the assertion in item (i) follows. Similarly, we can get (ii). □

Corollary 5.13. *For any $A \in \mathcal{V}(m,n)$, we have*

(i) $A \circ B \in \mathcal{S}(m,n)$, *for any* $B \in \mathcal{S}(m,n)$;

(ii) $A \circ B \in \mathcal{V}(m,n)$, *for any* $B \in \mathcal{V}(m,n)$.

Proof. Let $A = \sum_{i=1}^{s}[u^{(i)}]^m$. It follows from Proposition 2.6(iii) that

$$A \circ B = \sum_{i=1}^{s} (\bar{P}^{(i)})^{[m]}(B), \quad \forall B \in \mathcal{T}(m,n).$$

If $B \in \mathcal{S}(m,n)$, Theorem 5.12(i) indicates that

$$(\bar{P}^{(i)})^{[m]}(B) \in \mathcal{S}(m,n), \quad \forall\, i = 1,\ldots,s,$$

where $\bar{P}^{(i)}$ is the diagonal matrix generated by $u^{(i)}$. By the fact that $\mathcal{S}(m,n)$ is a convex cone, we immediately get that $A \circ B \in \mathcal{S}(m,n)$. Likely, we can obtain (ii). This completes the proof. □

When $m = 2$, the results in Corollary 5.13 turns out to be one of the basic results in matrix theory that the Hadamard product of two symmetric positive semidefinite matrices remains positive semidefinite; see [11, Theorem 5.2.1]. However, the Hadamard product of two symmetric semidefinite tensors of order $m \geqslant 4$ may fail to be positive semidefinite, as the following example shows.

Example 5.14. Let A, $B \in \mathcal{T}(4,n)$ with $A_{1122} = A_{1212} = A_{1221} = A_{2112} = A_{2121} = A_{2211} = 1$ and other entries of A zero, $B_{1111} = 1$, $B_{2222} = 9$, $B_{1122} = B_{1212} = B_{1221} = B_{2112} = B_{2121} = B_{2211} = -1$ and other entries of B zero. It is easy to verify that A, $B \in \mathcal{S}(4,n)$. However, $(A \circ B)[x]^4 = -6x_1^2x_2^2$ which implies that $A \circ B \notin \mathcal{S}(4,n)$.

Actually, the tensors A and B defined in the above example lie in $\mathrm{SOS}(4,n)$, which also shows that the Hadamard product of two tensors in $\mathrm{SOS}(m,n)$ may not be in $\mathrm{SOS}(m,n)$ anymore. Additionally, by summarizing Theorem 5.5(i) and Corollary 5.13(ii), we can easily get the following relationship to the nonnegative tensor cone.

Proposition 5.15. *Let $N(m,n)$ be the m-order n-dimensional nonnegative tensor cone. We have*

(i) $D_+(m,n) \subset \mathcal{S}(m,n) \cap N(m,n)$;

(ii) $\{A \circ A : A \in \mathcal{V}(m,n)\} \subset \mathcal{V}(m,n) \cap N(m,n)$.

The relationship among sets $\mathcal{S}(m,n)$, $\mathrm{SOS}(m,n)$, $\mathcal{V}(m,n)$, $D_+(m,n)$, $D_0(m,n)$, $N(m,n)$, and the set of all nonnegative diagonal tensors, denoted by $\bar{D}_0(m,n)$, are in Figure 1.

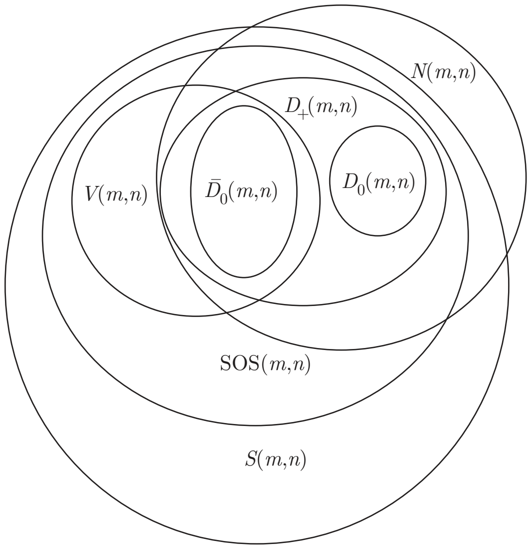

Figure 1 Relationship among different tensor cones

Specially, when $m = 2$, we have
- $\mathcal{S}(m,n) = \mathrm{SOS}(m,n) = \mathcal{V}(m,n)$;
- $D_+(m,n) = \bar{D}_+(m,n)$;
- $D_0(m,n) = \emptyset$.

6 Concluding remarks

We have employed an m-order generalization of symmetric-scaled representation used in interior-point algorithms in SDP contexts to linear operators over symmetric tensor spaces and have proved a decomposition invariance theorem for such a class of linear operators. This further leads to several other interesting properties in the underlying tensor spaces. We then considered the positive semidefiniteness of linear operators and managed to achieve the convexity of the Frobenius norm function of a symmetric tensor. Furthermore, by employing the properties of linear operators, we characterize the symmetric positive semidefinite tensor cone, design some face representations of its dual cone, and analyze its relations to many other convex cones, such as nonnegative tensor cone, partially diagonal tensor cone, and SOS cone. In particular, we show that the cone is self-dual if and only if the polynomial is quadratic, give specific characterizations of tensors that are in the primal cone but not in the dual, and develop a complete relationship map among the tensor cones appeared in the literature. Much more work needs to be done for the SDT cone both on theory analysis and numerical computations in the future.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 11301022), the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University (Grant Nos. RCS2014ZT20 and RCS2014ZZ001), Beijing Natural Science Foundation (Grant No. 9144031), and the Hong Kong Research Grant Council (Grant Nos. PolyU 501909, 502510, 502111 and 501212).

References

1 Ahmadi A A, Parrilo P A. Sum of squares and polynomial convexity. Preprint, 2009, http://www.researchgate.net/publication/228933473_Sum_of_Squares_and_Polynomial_Convexity

2 Alizadeh F. Interior point methods in semidefinite programming with applications to combinatorial optimization. SIAM J Optim, 1995, 5: 13–51

3 Barmpoutis A, Jian B, Vemuri B C, et al. Symmetric positive 4-th order tensors: Their estimation from diffusion weighted MRI. Inf Process Med Imaging, 2007, 20: 308–319

4 Bulò S R, Pelillo M. New bounds on the clique number of graphs based on spectral hypergraph theory. Lecture Notes Comput Sci, 2009, 5851: 45–58

5 Chang K C, Pearson K, Zhang T. Perron-Frobenius theorem for nonnegative tensors. Commun Math Sci, 2008, 6: 507–520

6 Comon P, Golub G, Lim L H, et al. Symmetric tensors and symmetric tensor rank. SIAM J Matrix Anal Appl, 2008, 30: 1254–1279

7 Comon P, Sorensen M. Tensor diagonalization by orthogonal transforms. Report ISRN I3S-RR-2007-06-FR, 2007

8 Faraut J, Korányi A. Analysis on Symmetric Cones. New York: Oxford University Press, 1994

9 He S, Li Z, Zhang S. Approximation algorithms for homogeneous polynomial optimization with quadratic constraints. Math Program, 2010, 125: 353–383

10 Helton J W, Nie J W. Semidefinite representation of convex sets. Math Program, 2010, 122: 21–64

11 Horn R A, Johnson C R. Topics in Matrix Analysis. Cambridge: Cambridge University Press, 1991

12 Kolda T G, Bader B W. Tensor decompositions and applications. SIAM Rev, 2009, 51: 455–500

13 Lang S. Real and Functional Analysis, 3rd ed. New York: Springer-Verlag, 1993

14 Lim L H. Multilinear pagerank: Measuring higher order connectivity in linked objects. The Internet: Today and Tomorrow, 2005

15 McCullagh P. Tensor Methods in Statistics. London: Chapman and Hall, 1987

16 Ng M, Qi L Q, Zhou G L. Finding the largest eigenvalue of a nonnegative tensor. SIAM J Matrix Anal Appl, 2009, 31: 1090–1099

17 Qi L Q. Eigenvalues of a real supersymmetric tensor. J Symbolic Comput, 2005, 40: 1302–1324

18 Qi L Q, Ye Y Y. Space tensor conic programming. Comput Optim Appl, 2014, 59: 307–319

19 Schmieta S H, Alizadeh F. Extension of primal-dual interior point algorithms to symmetric cones. Math Program, 2003, 96: 409–438

20 Sun D F, Sun J. Löwner's operator and spectral functions in Euclidean Jordan algebras. Math Oper Res, 2008, 33: 421–445

21 Ye Y Y. Interior Point Algorithms: Theory and Analysis. New York: John Wiley & Sons, 1997