Operators and Matrices
Volume 19, Number 4 (2025), 479–494
doi:10.7153/oam-2025-19-31

SOLVABILITY OF THE SYLVESTER TENSOR EQUATION

LI LIANG

(*Communicated by B. Jacob*)

Abstract. Exploring the Sylvester tensor equation not only enhances our understanding of complex systems in various fields but also furnishes us with powerful mathematical instruments to tackle practical challenges. In this study, relying on distinct inequality relationships among the indices that define the positions of elements, we carry out a comprehensive and profound investigation into the solvability of the Sylvester tensor equation by leveraging the theory of block tensors. Furthermore, via rigorous mathematical derivations, we uncover the analytical solution to this arduous equation and offer a thorough analysis of the dimensionality of the solution space.

1. Introduction

The Sylvester equation

$$AX + XB = C$$

where A, B, and C are square matrices, and X is the unknown matrix, was initially proposed by the British mathematician James Joseph Sylvester in 1884. Subsequently, this equation has been generalized to incorporate tensor operations, giving rise to the Sylvester tensor equation

$$\mathscr{X} \times_1 A^{(1)} + \mathscr{X} \times_2 A^{(2)} + \cdots + \mathscr{X} \times_m A^{(m)} = \mathscr{P}, \tag{1}$$

with $A^{(k)} \in \mathbb{R}^{n_k \times n_k}$, $\mathscr{P} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_m}$ and the k-mode product [6] "$\times_k$". The Sylvester tensor equation frequently emerges in the finite element [5], finite difference [2] or spectral method [9,10] discretization of a high-dimensional partial differential equation.

Exploring the solution of the Sylvester tensor equation not only enhances our understanding of intricate systems but also furnishes us with powerful mathematical instruments to tackle real-world challenges. In the early stages, the numerical solution of the Sylvester-like tensor equation was discussed using some well-established methods, such as the Krylov subspace method [1,7], the Schur decomposition [10], and the preconditioned iterative solvers [4]. Recently, several novel algorithms have been developed to obtain numerical solutions for the Sylvester-like tensor equation. In 2020, Li et al. presented multiple numerical algorithms for solving the discrete Lyapunov tensor equation, including the simple iterative method, the gradient-based iterative method, and the residual norm steepest descent iterative method [11]. In 2023, Zhang et al.

Mathematics subject classification (2020): 15A18, 15A69.
Keywords and phrases: Sylvester tensor equation, solvability, solution, solution space.

Paper OaM-19-31

proposed the tensor form of the generalized product-type biconjugate gradient method for solving generalized Sylvester quaternion tensor equations [15]. In the same year, Li et al. investigated the generalized coupled Sylvester tensor equations using the tensor forms of the biconjugate A-orthogonal residual and the conjugate A-orthogonal residual squared algorithms [8]. In 2024, Zhang et al. put forward the biconjugate A-orthogonal residual stabilized and the generalized product-type biconjugate A-orthogonal residual methods in tensor forms for solving the generalized Sylvester quaternion tensor equations [16]. In contrast, the analytical solution of the Sylvester-like tensor equation rarely emerges.

In this research, the solvability of Eq. (1) is established through the utilization of the Jordan canonical form and the block tensor. Subsequently, the analytical solution and the dimension of the solution space of the equation are presented. Specifically, the transformation of $A^{(k)} \in \mathbb{R}^{n_k \times n_k}$ to the Jordan canonical forms $J^{(k)} = U^{(k)} A^{(k)} (U^{(k)})^{-1}$ transforms Eq. (1) into

$$\mathscr{Y} \times_1 J^{(1)} + \mathscr{Y} \times_2 J^{(2)} + \cdots + \mathscr{Y} \times_m J^{(m)} = \mathscr{Q} \tag{2}$$

where $\mathscr{Y} = \mathscr{X} \times_1 U^{(1)} \times \cdots \times_m U^{(m)}$ and $\mathscr{Q} = \mathscr{P} \times_1 U^{(1)} \times \cdots \times_m U^{(m)}$. Then, it is decomposed into smaller equations based on the sub-blocks of $J^{(k)}$, $k = 1, 2, \ldots, m$. Evidently, the solvability, the analytical solution, and the dimension of the solution space of Eq. (2) are determined by those of all the smaller equations. Consequently, it is sufficient to conduct an investigation on these smaller equations.

The present paper is structured as follows. In Section 2, certain preliminary materials are introduced. In Section 3, the primary results are presented, encompassing the solvability, the analytical solution, and the dimension of the solution space of the Sylvester tensor equation. In Section 4, two numerical examples are provided. Finally, the main results are proven in the appendix.

2. Preliminaries

An $n_1 \times n_2 \times \cdots \times n_m$ tensor $\mathscr{A}$ is an array $[\mathscr{A}(l_1, l_2, \ldots, l_m)]_{l_1, l_2, \ldots, l_m = 1}^{n_1, n_2, \ldots, n_m}$, where $\mathscr{A}(l_1, l_2, \ldots, l_m)$ denotes the $(l_1, l_2, \ldots, l_m)$-th component of $\mathscr{A}$. The act of partitioning a tensor $\mathscr{A}$ into smaller tensors results in a block tensor $[\mathscr{A}_{i_1, i_2, \ldots, i_m}]$, where $\mathscr{A}_{i_1, i_2, \ldots, i_m}$ represents the $(i_1, i_2, \ldots, i_m)$-th block within $\mathscr{A}$. Evidently, $\mathscr{A}_{i_1, i_2, \ldots, i_m}(l_1, l_2, \ldots, l_m)$ is the $(l_1, l_2, \ldots, l_m)$-th component of the $(i_1, i_2, \ldots, i_m)$-th block in $\mathscr{A}$.

The question arises: how can a tensor $\mathscr{A}$ be partitioned into the block tensor $[\mathscr{A}_{i_1, i_2, \ldots, i_m}]$?

Partitioning an $n_1 \times n_2$ matrix A into a block tensor $[\mathscr{A}_{i_1, i_2}]_{i_1, i_2 = 1}^{b_1, b_2}$ involves dividing row and column vectors into block vectors

$$\mathbf{m}^{(1)} = [m_1^{(1)}, \ldots, m_{b_1}^{(1)}] \quad \text{and} \quad \mathbf{m}^{(2)} = [m_1^{(2)}, \ldots, m_{b_2}^{(2)}],$$

where $m_1^{(1)}, \ldots, m_{b_1}^{(1)}, m_1^{(2)}, \ldots, m_{b_2}^{(2)} \in \mathbb{N}$ and $m_1^{(1)} + \cdots + m_{b_1}^{(1)} = n_1$, $m_1^{(2)} + \cdots + m_{b_2}^{(2)} = n_2$. Analogously, an $n_1 \times n_2 \times \cdots \times n_m$ tensor $\mathscr{A}$ can be partitioned into a block tensor

[12] by defining the block vectors

$$\mathbf{m}^{(k)} = [m_1^{(k)}, \ldots, m_{b_k}^{(k)}], \quad k = 1, 2, \ldots, m,$$

with $m_1^{(k)}, \ldots, m_{b_k}^{(k)} \in \mathbb{N}$, $m_1^{(k)} + \cdots + m_{b_k}^{(k)} = n_k$. Indeed, the $(i_1, i_2, \ldots, i_m)$-th block in $\mathscr{A}$ is

$$\mathscr{A}_{i_1,i_2,\ldots,i_m} = \mathscr{A}(\rho_{i_1}^{(1)} + 1 : \rho_{i_1}^{(1)} + m_{i_1}^{(1)}, \ldots, \rho_{i_m}^{(m)} + 1 : \rho_{i_m}^{(m)} + m_{i_m}^{(m)}),$$

where $\rho_{i_k}^{(k)} = m_1^{(k)} + \cdots + m_{i_k-1}^{(k)}$ and $\rho_{i_k}^{(k)} + 1 : \rho_{i_k}^{(k)} + m_{i_k}^{(k)}$, $k = 1, 2, \ldots, m$, is the index range.

The block tensor finds application in the problem of blocked multilinear products [13]. Suppose there exists an $n_1 \times \cdots \times n_m$ tensor $\mathscr{A}$ and $q_k \times n_k$ matrices $B^{(k)}$ with block vectors $[\mathbf{m}_k^{(1)}, \mathbf{m}_k^{(2)}]$, $k = 1, 2, \ldots, m$. Let

$$\mathscr{C} \triangleq \mathscr{A} \times_1 B^{(1)} \times \cdots \times_m B^{(m)},$$

then the $(i_1, i_2, \ldots, i_m)$-th block in $\mathscr{C}$ is

$$\mathscr{C}_{i_1,i_2,\ldots,i_m} = \sum_{j_1,j_2,\ldots,j_m} \mathscr{A}_{j_1,j_2,\ldots,j_m} \times_1 B_{i_1,j_1}^{(1)} \times \cdots \times_m B_{i_m,j_m}^{(m)}.$$

3. Solvability of the Sylvester tensor equation

Let

$$J^{(k)} = \begin{pmatrix} (J^{(k)})_{11} & & & \\ & (J^{(k)})_{22} & & \\ & & \ddots & \\ & & & (J^{(k)})_{s_k,s_k} \end{pmatrix}_{n_k \times n_k},$$

$$(J^{(k)})_{i_j,i_j} = \lambda_{i_j}^{(k)} I_{t_{i_j}^k} + N_{t_{i_j}^k}, \quad N_{t_{i_j}^k} = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & & 1 \\ & & & & 0 \end{pmatrix}_{t_{i_j}^k \times t_{i_j}^k},$$

where $i_j = 1, 2, \ldots, s_k$, $t_1^k + t_2^k + \cdots + t_{s_k}^k = n_k$ and $k = 1, 2, \ldots, m$. Then, Eq. (2) can be partitioned into a series of smaller equations

$$\mathscr{Y}_{i_1,i_2,\ldots,i_m} \times_1 (J^{(1)})_{i_1,i_1} + \cdots + \mathscr{Y}_{i_1,i_2,\ldots,i_m} \times_m (J^{(m)})_{i_m,\ i_m} = \mathscr{Q}_{i_1,i_2,\ldots,i_m} \tag{3}$$

for $\mathscr{Y}_{i_1,i_2,\ldots,i_m} \in \mathbb{C}^{t_{i_1}^1 \times t_{i_2}^2 \times \cdots \times t_{i_m}^m}$, $i_k = 1, 2, \ldots, s_k$, $k = 1, 2, \ldots, m$, by blocking $\mathscr{Y}$ with the block matrix $J^{(k)}$, $k = 1, 2, \ldots, m$. Consequently, it is enough to consider the solvability of Eq. (3).

LEMMA 1. *The component of the solution of Eq.* (3) *is one of the following*

(i) $\mathscr{Y}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,l_m)$

$$=\frac{\mathscr{Q}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,l_m)-\sum_{j_1=1}^{t_{i_1}^1}\mathscr{Y}_{i_1,i_2,\ldots,i_m}(j_1,l_2,\ldots,l_m)N_{t_{i_1}^1}(l_1,j_1)-\cdots-\sum_{j_m=1}^{t_{i_m}^m}\mathscr{Y}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,j_m)N_{t_{i_m}^m}(l_m,j_m)}{\lambda_{i_1}^{(1)}+\cdots+\lambda_{i_m}^{(m)}}$$

if and only if $\lambda_{i_1}^{(1)}+\cdots+\lambda_{i_m}^{(m)}\neq 0$;

(ii) $\mathscr{Y}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,l_m)$ *is an arbitrary parameter if and only if* $\lambda_{i_1}^{(1)}+\cdots+\lambda_{i_m}^{(m)}=0$ *and*

$$\mathscr{Q}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,l_m)-\sum_{j_1=1}^{t_{i_1}^1}\mathscr{Y}_{i_1,i_2,\ldots,i_m}(j_1,l_2,\ldots,l_m)N_{t_{i_1}^1}(l_1,j_1)$$

$$-\cdots-\sum_{j_m=1}^{t_{i_m}^m}\mathscr{Y}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,j_m)N_{t_{i_m}^m}(l_m,j_m)=0;$$

(iii) $\mathscr{Y}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,l_m)=\emptyset$ *if and only if* $\lambda_{i_1}^{(1)}+\cdots+\lambda_{i_m}^{(m)}=0$ *and*

$$\mathscr{Q}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,l_m)-\sum_{j_1=1}^{t_{i_1}^1}\mathscr{Y}_{i_1,i_2,\ldots,i_m}(j_1,l_2,\ldots,l_m)N_{t_{i_1}^1}(l_1,j_1)$$

$$-\cdots-\sum_{j_m=1}^{t_{i_m}^m}\mathscr{Y}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,j_m)N_{t_{i_m}^m}(l_m,j_m)\neq 0.$$

Proof. Eq. (3) displays that the equation for $\mathscr{Y}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,l_m)$ is

$$(\lambda_{i_1}^{(1)}+\cdots+\lambda_{i_m}^{(m)})\mathscr{Y}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,l_m)=\mathscr{Q}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,l_m)$$

$$-\sum_{j_1=1}^{t_{i_1}^1}\mathscr{Y}_{i_1,i_2,\ldots,i_m}(j_1,l_2,\ldots,l_m)N_{t_{i_1}^1}(l_1,j_1)$$

$$-\cdots-\sum_{j_m=1}^{t_{i_m}^m}\mathscr{Y}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,j_m)N_{t_{i_m}^m}(l_m,j_m). \quad (4)$$

Hence, the result is obvious. □

It is evident that fulfilling either condition (i) or (ii) as stipulated in Lemma 1 results in the solvability and derivation of an analytical solution for Eq. (3). Additionally, the application of an iterative method facilitates the reformulation of conditions (i) and (ii), thereby yielding more expedient results.

THEOREM 1. *Assuming* $t_{i_q}^q = \max\{t_{i_1}^1, t_{i_2}^2, \ldots, t_{i_m}^m\}$,

(i) *if* $\lambda_{i_1}^{(1)} + \cdots + \lambda_{i_m}^{(m)} \neq 0$, *then the solution of Eq.* (3) *is existent and unique;*

(ii) *if* $\lambda_{i_1}^{(1)} + \cdots + \lambda_{i_m}^{(m)} = 0$, *then the solution of Eq.* (3) *is existent if and only if the components of* $\mathscr{Q}_{i_1,i_2,\ldots,i_m}$ *meet the requirements*

(a) *for all* $k_1, \ldots, k_{q-1}, k_{q+1}, \ldots, k_m$, *meeting* $k_j \in \{0, \ldots, t_{i_j}^j - 1\}$, $j = 1, \ldots, q-1, q+1, \ldots, m$, *and* $k_1 + \cdots + k_{q-1} + k_{q+1} + \cdots + k_m < t_{i_q}^q$,

$$\sum_{r=0}^{\sum_{i\in[m]\setminus\{q\}} k_i} (-1)^r \cdot \sum_{\substack{u_j = \max\{r - \sum_{i\in[j-1]\setminus\{q\}} u_i - \sum_{i\in[j+1,m]\setminus\{q\}} k_i, 0\} \\ j\in[m-1]\setminus\{q\}}}^{\min\{r - \sum_{i\in[j-1]\setminus\{q\}} u_i, k_j\}} \cdot \Big(\bigcup_j C_{r - \sum_{i\in[j-1]\setminus\{q\}} u_i}^{u_j}\Big) \mathscr{Q}_{i_1,i_2,\ldots,i_m}(w_1, \ldots, w_{q-1}, t_{i_q}^q - r, w_{q+1}, \ldots, w_m) = 0$$

(b) *for all* $k_1, \ldots, k_{q-1}, k_{q+1}, \ldots, k_m$, *meeting* $k_j \in \{0, 1, \ldots, t_{i_j}^j - 1\}$, $j = 1, \ldots, q-1, q+1, \ldots, m$, *and* $k_1 + \cdots + k_{q-1} + k_{q+1} + \cdots + k_m \geqslant t_{i_q}^q$, *solutions to the system of non-homogeneous linear equations*

$$\begin{aligned}
&\sum_{r=0}^{t_{i_q}^q - 1} (-1)^r \cdot \sum_{\substack{u_j = \max\{r - \sum_{i\in[j-1]\setminus\{q\}} u_i - \sum_{i\in[j+1,m]\setminus\{q\}} k_i, 0\} \\ j\in[m-1]\setminus\{q\}}}^{\min\{r - \sum_{i\in[j-1]\setminus\{q\}} u_i, k_j\}} \cdot \Big(\bigcup_j C_{r - \sum_{i\in[j-1]\setminus\{q\}} u_i}^{u_j}\Big) \mathscr{Q}_{i_1,i_2,\ldots,i_m}(w_1, \ldots, w_{q-1}, t_{i_q}^q - r, w_{q+1}, \ldots, w_m) \\
&+ (-1)^{t_{i_q}^q} \cdot \sum_{\substack{u_j = \max\{t_{i_q}^q - \sum_{i\in[j-1]\setminus\{q\}} u_i - \sum_{i\in[j+1,m]\setminus\{q\}} k_i, 0\} \\ j\in[m-1]\setminus\{q\}}}^{\min\{t_{i_q}^q - \sum_{i\in[j-1]\setminus\{q\}} u_i, k_j\}} \cdot \Big(\bigcup_j C_{t_{i_q}^q - \sum_{i\in[j-1]\setminus\{q\}} u_i}^{u_j}\Big) x(w_1, \ldots, w_{q-1}, 1, w_{q+1}, \ldots, \hat{w}_m) \\
&= 0
\end{aligned}$$

exist,

where $w_l = t_{i_l}^l - (k_l - u_l), \forall l \in [1, m-1]\setminus\{q\}$, $w_m = t_{i_m}^m - [k_m - (r - \sum_{l\in[m-1]\setminus\{q\}} u_l)]$ *and* $\hat{w}_m = t_{i_m}^m - [k_m - (t_{i_q}^q - \sum_{l\in[m-1]\setminus\{q\}} u_l)]$.

Theorem 1 is proved in the appendix.

THEOREM 2. *Assuming* $t_{i_q}^q = \max\{t_{i_1}^1, t_{i_2}^2, \ldots, t_{i_m}^m\}$ *and Eq.* (3) *is solvable, the solution of Eq.* (3) *is composed of the following two forms*

(i) *if* $\lambda_{i_1}^1 + \cdots + \lambda_{i_m}^m \neq 0$, *then*

$$\mathscr{Y}_{i_1,i_2,\ldots,i_m}$$

$$= \sum_{t=0}^{t_{i_1}^1+\cdots+t_{i_m}^m-m} (-1)^t (\lambda_{i_1}^1 + \cdots + \lambda_{i_m}^m)^{-(t+1)}$$

$$\cdot \Big(\sum_{\sigma_1+\cdots+\sigma_m=t} C_t^{\sigma_1} C_{t-\sigma_1}^{\sigma_2} \cdots C_{t-\sigma_1-\cdots-\sigma_{m-2}}^{\sigma_{m-1}} \mathscr{Q}_{i_1,i_2,\ldots,i_m} \times_1 N_{t_{i_1}^1}^{\sigma_1} \times \cdots \times_m N_{t_{i_m}^m}^{\sigma_m} \Big).$$

(ii) *if* $\lambda_{i_1}^1 + \cdots + \lambda_{i_m}^m = 0$, *then the* $(t_{i_1}^1 - k_1, \ldots, t_{i_m}^m - k_m)$*-th,* $k_j \in \{0, \ldots, t_{i_j}^j - 1\}$, $j = 1, 2, \ldots, m$, *component of* $\mathscr{Y}_{i_1,\ldots,i_m}$ *is*

(a) *when* $k_1 + \cdots + k_{q-1} + k_{q+1} + \cdots + k_m < t_{i_q}^q - 1 - k_q$,

$$\mathscr{Y}_{i_1,i_2,\ldots,i_m}(t_{i_1}^1 - k_1, \ldots, t_{i_m}^m - k_m)$$

$$= \sum_{r=0}^{\sum_{i\in[m]\setminus\{q\}} k_i} (-1)^r \cdot \sum_{\substack{u_j = \max\{r - \sum_{i\in[j-1]\setminus\{q\}} u_i - \sum_{i\in[j+1,m]\setminus\{q\}} k_i, 0\} \\ j\in[m-1]\setminus\{q\}}}^{\min\{r-\sum_{i\in[j-1]\setminus\{q\}} u_i, k_j\}}$$

$$\cdot \Big(\bigcup_j C_{r - \sum_{i\in[j-1]\setminus\{q\}} u_i}^{u_j} \Big) \mathscr{Q}_{i_1,i_2,\ldots,i_m}(w_1, \ldots, w_{q-1}, t_{i_q}^q - (r+1) - k_q, w_{q+1}, \ldots, w_m)$$

(b) *when* $k_1 + \cdots + k_{q-1} + k_{q+1} + \cdots + k_m \geqslant t_{i_q}^q - 1 - k_q$,

$$\mathscr{Y}_{i_1,i_2,\ldots,i_m}(t_{i_1}^1 - k_1, \ldots, t_{i_m}^m - k_m)$$

$$= \sum_{r=0}^{t_{i_q}^q - 1 - (k_q+1)} (-1)^r \cdot \sum_{\substack{u_j = \max\{r - \sum_{i\in[j-1]\setminus\{q\}} u_i - \sum_{i\in[j+1,m]\setminus\{q\}} k_i, 0\} \\ j\in[m-1]\setminus\{q\}}}^{\min\{r-\sum_{i\in[j-1]\setminus\{q\}} u_i, k_j\}}$$

$$\cdot \Big(\bigcup_j C_{r - \sum_{i\in[j-1]\setminus\{q\}} u_i}^{u_j} \Big) \mathscr{Q}_{i_1,i_2,\ldots,i_m}(w_1, \ldots, w_{q-1}, t_{i_q}^q - (r+1) - k_q, w_{q+1}, \ldots, w_m)$$

$$+ (-1)^{t_{i_q}^q - (k_q+1)} \cdot \sum_{\substack{u_j = \max\{t_{i_q}^q - (k_q+1) - \sum_{i\in[j-1]\setminus\{q\}} u_i - \sum_{i\in[j+1,m]\setminus\{q\}} k_i, 0\} \\ j\in[m-1]\setminus\{q\}}}^{\min\{t_{i_q}^q - (k_q+1) - \sum_{i\in[j-1]\setminus\{q\}} u_i, k_j\}}$$

$$\cdot \Big(\bigcup_j C_{t_{i_q}^q - (k_q+1) - \sum_{i\in[j-1]\setminus\{q\}} u_i}^{u_j} \Big) \mathscr{Y}_{i_1,i_2,\ldots,i_m}(w_1, \ldots, w_{q-1}, 1, w_{q+1}, \ldots, \hat{w}_m)$$

where where $w_l = t^l_{i_l} - (k_l - u_l), \forall l \in [m-1]\backslash\{q\}$, $w_m = t^m_{i_m} - [k_m - (r - \sum_{l\in[m-1]\backslash\{q\}} u_l)]$ *and* $\hat{w}_m = t^m_{i_m} - [k_m - (t^q_{i_q} - (k_q + 1) - \sum_{l\in[m-1]\backslash\{q\}} u_l)]$.

Significantly, if the indexes in the parenthesis of components

$$\mathscr{Y}_{i_1,i_2,\ldots,i_m}(t^1_{i_1} - k_1, \ldots, t^{q-1}_{i_{q-1}} - k_{q-1}, 1, t^{q+1}_{i_{q+1}} - k_{q+1}, \ldots, t^m_{i_m} - k_m)$$

are the same as that of the variate x of the system in Theorem 1(ii)(b), then they satisfy the system; otherwise are arbitrary parameters.

Theorem 2 is proved in the appendix.

Under the solvability condition, we now proceed to the exploration of the dimension of the solution space of Eq. (3) for the undetermined values of the solution in Theorem 2.

THEOREM 3. *Assuming* $t^q_{i_q} = \max\{t^1_{i_1}, t^2_{i_2}, \ldots, t^m_{i_m}\}$ *and Eq.* (3) *is solvable, the dimension of the solution space of Eq.* (3) *is one of the following*

(i) *if* $\lambda^1_{i_1} + \cdots + \lambda^m_{i_m} \neq 0$, *the dimension of the solution space is 0;*

(ii) *if* $\lambda^1_{i_1} + \cdots + \lambda^m_{i_m} = 0$, *the dimension of the solution space is*

$$t^1_{i_1} \cdots t^{q-1}_{i_{q-1}} t^{q+1}_{i_{q+1}} \cdots t^m_{i_m} - r$$

where r *is the rank of the coefficient matrix of the system in Theorem 1(ii)(b), and*

(a) *if* $t^1_{i_1} + \cdots + t^{q-1}_{i_{q-1}} + t^{q+1}_{i_{q+1}} + \cdots + t^m_{i_m} - (m-1) < t^q_{i_q}$, *then* $r = 0$;

(b) *if* $t^1_{i_1} + \cdots + t^{q-1}_{i_{q-1}} + t^{q+1}_{i_{q+1}} + \cdots + t^m_{i_m} - (m-1) \geqslant t^q_{i_q}$, *then*

$$1 \leqslant r \leqslant t^1_{i_1} \cdots t^{q-1}_{i_{q-1}} t^{q+1}_{i_{q+1}} \cdots t^m_{i_m} - \sum_{n=0}^{t^q_{i_q}-1} [C^{m-2}_{m+n-2} + \sum_{t=1}^{m-1} (-1)^t \sum_{\substack{j_1<\cdots<j_t \\ j_1,\ldots,j_t\in[m]\backslash\{q\}}} C^{m-2}_{m+n-2-\sum_{l=1}^t t^{j_l}_{i_{j_l}}}],$$

where $C^{m-2}_{m+n-2-\sum_{l=1}^t t^{j_l}_{i_{j_l}}}$ *is valid when* $n \geqslant \sum_{l=1}^t t^{j_l}_{i_{j_l}}$.

Theorem 3 is proved in the appendix.

4. Numerical example

In this section, we substantiate the aforementioned results by presenting two illustrative examples.

EXAMPLE 1. Determine the solution to the equation

$$\mathscr{X} \times_1 A_1 + \mathscr{X} \times_2 A_2 + \mathscr{X} \times_3 A_3 = \mathscr{P},$$

where matrices A_1, A_2 and A_3

$$A_1 = \begin{pmatrix} 3 & -4 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 4 & 3 & 4 & 1 \\ -4 & -7 & -4 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} -2 & 1 \\ 0 & -2 \end{pmatrix}$$

and slices $\mathscr{P}(:: 1)$, $\mathscr{P}(:: 2)$ of $\mathscr{P}$

$$\mathscr{P}(:: 1) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \quad \mathscr{P}(:: 2) = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

According to the Jordan canonical forms $J_j = U_j A_j U_j^{-1}$ of the matrices A_j, $j = 1,2,3$, the original equation is transformed into

$$\mathscr{Y} \times_1 J_1 + \mathscr{Y} \times_2 J_2 + \mathscr{Y} \times_3 J_3 = \mathscr{Q}$$

with

$$\mathscr{Y} = \mathscr{X} \times_1 U_1 \times_2 U_2 \times_3 U_3, \ \mathscr{Q} = \mathscr{P} \times_1 U_1 \times_2 U_2 \times_3 U_3.$$

Now we partition $\mathscr{Y}$ and $\mathscr{Q}$ into smaller tensors $\mathscr{Y}_{i_1,i_2,i_3}$ and $\mathscr{Q}_{i_1,i_2,i_3}$, $i_1 = 1,2$, $i_2 = 1,2$, $i_3 = 1$, by

$$J_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \quad J_3 = \begin{pmatrix} -2 & 1 \\ 0 & -2 \end{pmatrix},$$

According to Theorem 1 and 2, the solutions of the above block tensor equations as follows:

$$\mathscr{Y}_{1,1,1}(:: 1) = \begin{pmatrix} y_{111} & y_{121} \\ y_{211} & \frac{1}{2} \end{pmatrix}, \quad \mathscr{Y}_{1,1,1}(:: 2) = \begin{pmatrix} -(y_{121} + y_{211}) & -\frac{3}{2} \\ \frac{1}{2} & -1 \end{pmatrix}$$

with three arbitrary parameters y_{111}, y_{121}, y_{211};

$$\mathscr{Y}_{1,2,1}(:: 1) = \begin{pmatrix} 4 & -2 \\ -6 & 5 \end{pmatrix}, \quad \mathscr{Y}_{1,2,1}(:: 2) = \begin{pmatrix} 4 & -6 \\ 2 & -1 \end{pmatrix};$$

$$\mathscr{Y}_{2,1,1}(:: 1) = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \mathscr{Y}_{2,1,1}(:: 2) = \begin{pmatrix} 0 & 1 \end{pmatrix};$$
$$\mathscr{Y}_{2,2,1}(:: 1) = \begin{pmatrix} \frac{1}{4} & \frac{5}{2} \end{pmatrix}, \quad \mathscr{Y}_{2,2,1}(:: 2) = \begin{pmatrix} -2 & 5 \end{pmatrix}.$$

Consequently, the solution of the original equation is

$$\mathscr{X}(:: 1) = \begin{pmatrix} -2y_{111} + y_{211} & -2y_{121} + \frac{1}{2} & -14 & 8y_{111} - 4y_{211} + 6y_{121} + \frac{71}{2} \\ -y_{111} + y_{211} & -y_{121} + \frac{1}{2} & -10 & 4y_{111} - 4y_{211} + 3y_{121} + \frac{51}{2} \\ 1 & 0 & \frac{1}{4} & -2 \end{pmatrix},$$

$$\mathscr{X}(:: 2) = \begin{pmatrix} 2(y_{121} + y_{211}) + \frac{1}{2} & 2 & -6 & -8(y_{121} + y_{211}) + 15 \\ y_{121} + y_{211} + \frac{1}{2} & \frac{1}{2} & -2 & -4(y_{121} + y_{211}) - \frac{11}{2} \\ 0 & 1 & -2 & 6 \end{pmatrix}.$$

EXAMPLE 2. Solve the equation

$$\mathscr{X} \times_1 A_1 + \mathscr{X} \times_2 A_2 + \mathscr{X} \times_3 A_3 + \mathscr{X} \times_4 A_4 = \mathscr{P}$$

with

$$A_1 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -3 & 1 \\ 0 & -3 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 3 & 6 \\ 0 & -3 \end{pmatrix}, \quad A_4 = \begin{pmatrix} -1 & 1 & 2 \\ 3 & 1 & -2 \\ -5 & 1 & 5 \end{pmatrix}$$

and slices of $\mathscr{P}$

$$\mathscr{P}(1\,1\,::) = \begin{pmatrix} 3 & 0 & 3 \\ -2 & 1 & -2 \end{pmatrix}, \qquad \mathscr{P}(1\,2\,::) = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix},$$

$$\mathscr{P}(2\,1\,::) = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & 0 \end{pmatrix}, \qquad \mathscr{P}(2\,2\,::) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

According to Jordan canonical forms $J_j = U_j A_j U_j^{-1}$ of matrices A_j, $j = 1,2,3,4$, the original equation is transformed into

$$\mathscr{Y} \times_1 J_1 + \mathscr{Y} \times_2 J_2 + \mathscr{Y} \times_3 J_3 + \mathscr{Y} \times_4 J_4 = \mathscr{Q}$$

with

$$\mathscr{Y} = \mathscr{X} \times_1 U_1 \times_2 U_2 \times_3 U_3 \times_4 U_4, \quad \mathscr{Q} = \mathscr{P} \times_1 U_1 \times_2 U_2 \times_3 U_3 \times_4 U_4,$$

where

$$\mathscr{Q}(1\,1\,::) = \begin{pmatrix} -1 & -1 & 2 \\ -1 & -1 & -1 \end{pmatrix}, \qquad \mathscr{Q}(1\,2\,::) = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 4 & 0 \end{pmatrix},$$

$$\mathscr{Q}(2\,1\,::) = \begin{pmatrix} -2 & -1 & 6 \\ -1 & 0 & 3 \end{pmatrix}, \qquad \mathscr{Q}(2\,2\,::) = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix};$$

Then $\mathscr{Y}$ and $\mathscr{Q}$ can be partitioned into smaller tensors $\mathscr{Y}_{i_1,i_2,i_3,i_4}$ and $\mathscr{Q}_{i_1,i_2,i_3,i_4}$, $i_1=1,2,\ i_2=1,\ i_3=1,2,\ i_4=1,2$ by

$$J_1=A_1=\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix},\ J_2=A_2=\begin{pmatrix}-3 & 1\\ 0 & -3\end{pmatrix},\ J_3=\begin{pmatrix}3 & 0\\ 0 & -3\end{pmatrix},\ J_4=\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 1\\ 0 & 0 & 2\end{pmatrix}.$$

According to Theorems 1 and 2, the solutions of the above block tensor equations are obtained, and they are combined to present the solution of the equation $\mathscr{Y}\times_1 J_1+\mathscr{Y}\times_2 J_2+\mathscr{Y}\times_3 J_3+\mathscr{Y}\times_4 J_4=\mathscr{Q}$ as follows:

$$\mathscr{Y}(1\ 1\ ::)=\begin{pmatrix}y_{1111} & -4 & 2\\ \frac{1}{9} & \frac{2}{25} & \frac{1}{5}\end{pmatrix},\qquad \mathscr{Y}(1\ 2\ ::)=\begin{pmatrix}-1 & 1 & 0\\ -\frac{1}{3} & -\frac{4}{5} & 0\end{pmatrix},$$

$$\mathscr{Y}(2\ 1\ ::)=\begin{pmatrix}y_{2111} & -4 & 4\\ \frac{1}{6} & -\frac{22}{125} & -\frac{16}{25}\end{pmatrix},\quad \mathscr{Y}(2\ 2\ ::)=\begin{pmatrix}-2 & -1 & 2\\ 0 & -\frac{6}{25} & -\frac{1}{5}\end{pmatrix};$$

with two arbitrary parameters $y_{1111},\ y_{2111}$.

Finally, the solution of the equation $\mathscr{X}\times_1 A_1+\mathscr{X}\times_2 A_2+\mathscr{X}\times_3 A_3+\mathscr{X}\times_4 A_4=\mathscr{P}$ by $\mathscr{X}=\mathscr{Y}\times_1 (U_1)^{-1}\times_2 (U_2)^{-1}\times_3 (U_3)^{-1}\times_4 (U_4)^{-1}$.

Appendix

Proof of Theorem 1

Proof. We just prove (ii) since (i) is well known [4, 14].

The conditions that $\lambda_{i_1}^{(1)}+\cdots+\lambda_{i_m}^{(m)}=0$ and the solvability change Eq. (4) into

$$\begin{aligned}&\mathscr{Q}_{i_1,i_2,\ldots,i_m}(l_1,l_2,\ldots,l_m)\\ &=\mathscr{Y}_{i_1,i_2,\ldots,i_m}(l_1+1,l_2,\ldots,l_m)+\cdots+\mathscr{Y}_{i_1,i_2,\ldots,i_m}(l_1,\ldots,l_{m-1},l_m+1)\end{aligned}\tag{5}$$

with $l_j=1,2,\ldots,t_{i_j}^j,\ j=1,2,\ldots,m$. Specially, the term $\mathscr{Y}_{i_1,i_2,\ldots,i_m}(l_1,\ldots,l_{j-1},l_j+1,l_{j+1},\ldots,l_m)$ is nonexistent when $l_j=t_{i_j}^j$.

Given $r,\ k_1,\ldots,k_{q-1},k_{q+1},\ldots,k_m$, we let $l_1=t_{i_1}^1-(k_1-u_1),\ldots,l_{q-1}=t_{i_{q-1}}^{q-1}-(k_{q-1}-u_{q-1}),\ l_q=t_{i_q}^q-r,\ l_{q+1}=t_{i_{q+1}}^{q+1}-(k_{q+1}-u_{q+1}),\ldots,l_{m-1}=t_{i_{m-1}}^{m-1}-(k_{m-1}-u_{m-1}),\ l_m=t_{i_m}^m-[k_m-(r-u_1-\cdots-u_{q-1}-u_{q+1}-\cdots-u_{m-1})]$ with $u_j=\max\{r-k_{j+1}-\cdots-k_{q-1}-k_{q+1}-\cdots-k_m-u_1-\cdots-u_{q-1}-u_{q+1}-\cdots-u_{j-1},0\},\ldots,\min\{r-u_1-\cdots-u_{q-1}-u_{q+1}-\cdots-u_{j-1},k_j\},\ j=1,\ldots,q-1,q+1,\ldots,m$, in Eq. (5), then

$$\sum_{\substack{u_j=\max\{r-\sum_{i\in[j-1]\setminus\{q\}}u_i-\sum_{i\in[j+1,m]\setminus\{q\}}k_i,0\}\\ j\in[m-1]\setminus\{q\}}}^{\min\{r-\sum_{i\in[j-1]\setminus\{q\}}u_i,k_j\}}\Big(\bigcup_j C_{r-\sum_{i\in[j-1]\setminus\{q\}}u_i}^{u_j}\Big)$$

$$\cdot\mathscr{Q}_{i_1,i_2,\ldots,i_m}(w_1,\ldots,w_{q-1},t_{i_q}^q-r,w_{q+1},\ldots,w_m)$$

$$= \sum_{\substack{u_j=\max\{r+1-\sum_{i\in[j-1]\setminus\{q\}}u_i-\sum_{i\in[j+1,m]\setminus\{q\}}k_i,0\}\\ j\in[m-1]\setminus\{q\}}}^{\min\{r+1-\sum_{i\in[j-1]\setminus\{q\}}u_i,k_j\}} (\bigcup_j C_{r+1-\sum_{i\in[j-1]\setminus\{q\}}u_i}^{u_j})$$

$$\cdot \mathscr{Y}_{i_1,i_2,\ldots,i_m}(w_1,\ldots,w_{q-1},t_{i_q}^q-r,w_{q+1},\ldots,\widetilde{w}_m)$$

$$+ \sum_{\substack{u_j=\max\{r-\sum_{i\in[j-1]\setminus\{q\}}u_i-\sum_{i\in[j+1,m]\setminus\{q\}}k_i,0\}\\ j\in[m-1]\setminus\{q\}}}^{\min\{r-\sum_{i\in[j-1]\setminus\{q\}}u_i,k_j\}} (\bigcup_j C_{r-\sum_{i\in[j-1]\setminus\{q\}}u_i}^{u_j})$$

$$\cdot \mathscr{Y}_{i_1,i_2,\ldots,i_m}(w_1,\ldots,w_{q-1},t_{i_q}^q-(r-1),w_{q+1},\ldots,w_m)$$

where $w_l = t_{i_l}^l-(k_l-u_l), \forall l\in[1,m-1]\setminus\{q\}$, $w_m = t_{i_m}^m-[k_m-(r-\sum_{l\in[1,m-1]\setminus\{q\}}u_l)]$ and $\widetilde{w}_m = t_{i_m}^m-[k_m-(r+1-\sum_{l\in[1,m-1]\setminus\{q\}}u_l)]$.

Summing all the equations with $r=0,1,\ldots,\min\{k_1+\cdots+k_{q-1}+k_{q+1}+\cdots+k_m,t_{i_q}^q-1\}$ yields Theorem 1(ii)(a)(b).

Now we prove the converse of Theorem 1(ii).

We will substitute $\mathscr{Y}_{i_1,i_2,\ldots,i_m}(t_{i_1}^1-k_1,\ldots,t_{i_{q-1}}^{q-1}-k_{q-1},t_{i_q}^q-k_q,t_{i_{q+1}}^{q+1}-k_{q+1},\ldots,t_{i_m}^m-k_m)$, $k_j=0,1,\ldots,t_{i_j}^j-1$, $j=1,2,\ldots,m$, into the left side of Eq. (3)

$$\mathscr{Y}_{i_1,i_2,\ldots,i_m}(t_{i_1}^1-(k_1-1),\ldots,t_{i_{q-1}}^{q-1}-k_{q-1},t_{i_q}^q-k_q,t_{i_{q+1}}^{q+1}-k_{q+1},\ldots,t_{i_m}^m-k_m)$$

$$+\mathscr{Y}_{i_1,i_2,\ldots,i_m}(t_{i_1}^1-k_1,t_{i_2}^2-(k_2-1),\ldots,t_{i_{q-1}}^{q-1}-k_{q-1},t_{i_q}^q-k_q,t_{i_{q+1}}^{q+1}-k_{q+1},\ldots,t_{i_m}^m-k_m)$$

$$+\cdots+\mathscr{Y}_{i_1,i_2,\ldots,i_m}(t_{i_1}^1-k_1,\ldots,t_{i_{q-1}}^{q-1}-k_{q-1},t_{i_q}^q-k_q,t_{i_{q+1}}^{q+1}-k_{q+1},\ldots,t_{i_m}^m-(k_m-1)),$$

to verify whether they are the solutions or not, where the term $\mathscr{Y}_{i_1,i_2,\ldots,i_m}(t_{i_1}^1-k_1,\ldots,t_{i_{j-1}}^{j-1}-k_{j-1},t_{i_j}^j-(k_j-1),t_{i_{j+1}}^{j+1}-k_{j+1},\ldots,t_{i_m}^m-k_m)$ is nonexistent when $k_j=0$.

1) When $k_1+\cdots+k_{q-1}+k_{q+1}+\cdots+k_m-1<t_{i_q}^q-1-k_q$

(a) and $k_q\neq 0$, the $(t_{i_1}^1-k_1,\ldots,t_{i_m}^m-k_m)$-th component in the left side of Eq. (3) is

$$\sum_{r=0}^{\sum_{i\in[m]\setminus\{q\}}k_i}(-1)^r\cdot \sum_{\substack{u_j=\max\{r+1-\sum_{i\in[j-1]\setminus\{q\}}u_i-\sum_{i\in[j+1,m]\setminus\{q\}}k_i,0\}\\ j\in[m-1]\setminus\{q\}}}^{\min\{r+1-\sum_{i\in[j-1]\setminus\{q\}}u_i,k_j\}}$$

$$\cdot(\bigcup C_{r+1-\sum_{i\in[1,j-1]\setminus\{q\}}u_i}^{u_j})\mathscr{Q}_{i_1,i_2,\ldots,i_m}(w_1,\ldots,w_{q-1},t_{i_q}^q-(r+1)-k_q,w_{q+1},\ldots,\widetilde{w}_m)$$

$$+\sum_{r=0}^{\sum_{i\in[m]\setminus\{q\}}k_i}(-1)^r\cdot \sum_{\substack{u_j=\max\{r-\sum_{i\in[j-1]\setminus\{q\}}u_i-\sum_{i\in[j+1,m]\setminus\{q\}}k_i,0\}\\ j\in[m-1]\setminus\{q\}}}^{\min\{r-\sum_{i\in[j-1]\setminus\{q\}}u_i,k_j\}}$$

$$\cdot\Big(\bigcup C^{u_j}_{r-\sum\limits_{i\in[j-1]\setminus\{q\}}u_i}\mathscr{Q}_{i_1,i_2,\ldots,i_m}(w_1,\ldots,w_{q-1},t^q_{i_q}-r-k_q,w_{q+1},\ldots,w_m)$$

$$+(-1)^0\mathscr{Q}_{i_1,i_2,\ldots,i_m}(t^1_{i_1}-k_1,\ldots,t^m_{i_m}-k_m)$$

$$=\mathscr{Q}_{i_1,i_2,\ldots,i_m}(t^1_{i_1}-k_1,\ldots,t^m_{i_m}-k_m)$$

(b) and $k_q=0$, the $(t^1_{i_1}-k_1,\ldots,t^m_{i_m}-k_m)$-th component in the left side of Eq. (3) is

$$\sum_{r=0}^{\sum_{i\in[m]\setminus\{q\}}k_i}(-1)^r\cdot\sum_{\substack{u_j=\max\{r+1-\sum_{i\in[j-1]\setminus\{q\}}u_i-\sum_{i\in[j+1,m]\setminus\{q\}}k_i,0\}\\ j\in[m-1]\setminus\{q\}}}^{\min\{r+1-\sum_{i\in[j-1]\setminus\{q\}}u_i,k_j\}}$$

$$\cdot\Big(\bigcup C^{u_j}_{r+1-\sum\limits_{i\in[j-1]\setminus\{q\}}u_i}\Big)\mathscr{Q}_{i_1,i_2,\ldots,i_m}(w_1,\ldots,w_{q-1},t^q_{i_q}-(r+1),w_{q+1},\ldots,\widetilde{w}_m)$$

According to Theorem 1(ii)(a), the left side of Eq. (3) is equal to

$$\mathscr{Q}_{i_1,i_2,\ldots,i_m}(t^1_{i_1}-k_1,\ldots,t^{q-1}_{i_{q-1}}-k_{q-1},t^q_{i_q},t^{q+1}_{i_{q+1}}-k_{q+1},\ldots,t^m_{i_m}-k_m)$$

where $w_l=t^l_{i_l}-(k_l-u_l),\forall l\in[m-1]\setminus\{q\}$, $w_m=t^m_{i_m}-[k_m-(r-\sum\limits_{l\in[1,m-1]\setminus\{q\}}u_l)]$ and $\widetilde{w}_m=t^m_{i_m}-[k_m-(r+1-\sum\limits_{l\in[m-1]\setminus\{q\}}u_l)]$.

2) When $k_1+\cdots+k_{q-1}+k_{q+1}+\cdots+k_m-1\geqslant t^q_{i_q}-1-k_q$

(a) and $k_q\neq 0$, the $(t^1_{i_1}-k_1,\ldots,t^m_{i_m}-k_m)$-th component in the left side of Eq. (3) is

$$\sum_{r=0}^{t^q_{i_q}-k_q-2}(-1)^r\cdot\sum_{\substack{u_j=\max\{r+1-\sum_{i\in[j-1]\setminus\{q\}}u_i-\sum_{i\in[j+1,m]\setminus\{q\}}k_i,0\}\\ j\in[m-1]\setminus\{q\}}}^{\min\{r+1-\sum_{i\in[j-1]\setminus\{q\}}u_i,k_j\}}$$

$$\cdot\Big(\bigcup_j[C^{u_j}_{r+1-\sum\limits_{i\in[j-1]\setminus\{q\}}u_i}\Big)\mathscr{Q}_{i_1,i_2,\ldots,i_m}(w_1,\ldots,w_{q-1},t^q_{i_q}-(r+1)-k_q,w_{q+1},\ldots,\widetilde{w}_m)$$

$$+\sum_{r=1}^{t^q_{i_q}-k_q-1}(-1)^r\cdot\sum_{\substack{u_j=\max\{r-\sum_{i\in[j-1]\setminus\{q\}}u_i-\sum_{i\in[j+1,m]\setminus\{q\}}k_i,0\}\\ j\in[m-1]\setminus\{q\}}}^{\min\{r-\sum_{i\in[j-1]\setminus\{q\}}u_i,k_j\}}$$

$$\cdot\Big(\bigcup[C^{u_j}_{r-\sum\limits_{i\in[j-1]\setminus\{q\}}u_i}\Big)\mathscr{Q}_{i_1,i_2,\ldots,i_m}(w_1,\ldots,w_{q-1},t^q_{i_q}-r-k_q,w_{q+1},\ldots,w_m)$$

$$+(-1)^0\mathscr{Q}_{i_1,i_2,\ldots,i_m}(t^1_{i_1}-k_1,\ldots,t^{q-1}_{i_{q-1}}-k_{q-1},t^q_{i_q}-k_q,t^{q+1}_{i_{q+1}}-k_{q+1},\ldots,t^m_{i_m}-k_m)$$

$$=\mathscr{Q}_{i_1,i_2,\ldots,i_m}(t^1_{i_1}-k_1,\ldots,t^{q-1}_{i_{q-1}}-k_{q-1},t^q_{i_q}-k_q,t^{q+1}_{i_{q+1}}-k_{q+1},\ldots,t^m_{i_m}-k_m).$$

(b) and $k_q = 0$, the $(t^1_{i_1} - k_1, \ldots, t^m_{i_m} - k_m)$-th component in the left side of Eq. (3) is

$$\sum_{r=0}^{t^q_{i_q}-2} (-1)^r \cdot \sum_{\substack{u_j = \max\{r+1-\sum_{i\in[j-1]\setminus\{q\}} u_i - \sum_{i\in[j+1,m]\setminus\{q\}} k_i, 0\} \\ j\in[m-1]\setminus\{q\}}}^{\min\{r+1-\sum_{i\in[j-1]\setminus\{q\}} u_i, k_j\}}$$

$$\cdot \Big(\bigcup_j C^{u_j}_{r+1-\sum_{i\in[j-1]\setminus\{q\}} u_i}\Big) \mathscr{Q}_{i_1,i_2,\ldots,i_m}(w_1,\ldots,w_{q-1}, t^q_{i_q} - (r+1), w_{q+1},\ldots,\widetilde{w}_m)$$

$$+(-1)^{t^q_{i_q}-1} \cdot \sum_{\substack{u_j = \max\{t^q_{i_q}-1-\sum_{i\in[j-1]\setminus\{q\}} u_i - \sum_{i\in[j+1,m]\setminus\{q\}} k_i, 0\} \\ j\in[m-1]\setminus\{q\}}}^{\min\{t^q_{i_q}-1-\sum_{i\in[j-1]\setminus\{q\}} u_i, k_j\}}$$

$$\cdot \Big(\bigcup_j C^{u_j}_{t^q_{i_q}-1-\sum_{i\in[j-1]\setminus\{q\}} u_i}\Big) \mathscr{Y}_{i_1,i_2,\ldots,i_m}(w_1,\ldots,w_{q-1},1,w_{q+1},\ldots,\hat{w}_m)$$

where $\hat{w}_m = t^m_{i_m} - [k_m - (t^q_{i_q} - \sum_{l\in[1,m-1]\setminus\{q\}} u_l)]$. Significantly, components

$$\mathscr{Y}_{i_1,i_2,\ldots,i_m}(w_1,\ldots,w_{q-1},1,w_{q+1},\ldots,w_{m-1},\hat{w}_m)$$

satisfy the system of the non-homogeneous linear equations in Theorem 1(ii)(b). Therefore, the left side of Eq. (3) is equal to

$$\mathscr{Q}_{i_1,i_2,\ldots,i_m}(t^1_{i_1} - k_1, \ldots, t^{q-1}_{i_{q-1}} - k_{q-1}, t^q_{i_q}, t^{q+1}_{i_{q+1}} - k_{q+1}, \ldots, t^m_{i_m} - k_m). \quad \square$$

Proof of Theorem 2

Proof.

(i) Eq. (3) can be written as

$$(\lambda^{(1)}_{i_1} + \cdots + \lambda^{(m)}_{i_m})\mathscr{Y}_{i_1,i_2,\ldots,i_m} = \mathscr{Q}_{i_1,i_2,\ldots,i_m} - \mathscr{Y}_{i_1,i_2,\ldots,i_m} \times_1 N_{t^1_{i_1}} - \cdots - \mathscr{Y}_{i_1,i_2,\ldots,i_m} \times_m N_{t^m_{i_m}}.$$

Multiplying $\lambda^{(1)}_{i_1} + \cdots + \lambda^{(m)}_{i_m}$ on both sides repeatedly gets

$$(\lambda^{(1)}_{i_1} + \cdots + \lambda^{(m)}_{i_m})^r \mathscr{Y}_{i_1,i_2,\ldots,i_m}$$

$$= \sum_{t=0}^{r-1} (-1)^t (\lambda^1_{i_1} + \cdots + \lambda^m_{i_m})^{r-(t+1)}$$

$$\cdot \sum_{\sigma_1+\cdots+\sigma_m=t} C^{\sigma_1}_t C^{\sigma_2}_{t-\sigma_1} \cdots C^{\sigma_{m-1}}_{t-\sigma_1-\cdots-\sigma_{m-2}} \mathscr{Q}_{i_1,i_2,\ldots,i_m} \times_1 N^{\sigma_1}_{t^1_{i_1}} \times \cdots \times_m N^{\sigma_m}_{t^m_{i_m}}$$

$$+(-1)^r \sum_{\sigma_1+\cdots+\sigma_m=r} C^{\sigma_1}_r C^{\sigma_2}_{r-\sigma_1} \cdots C^{\sigma_{m-1}}_{r-\sigma_1-\cdots-\sigma_{m-2}} \mathscr{Y}_{i_1,i_2,\ldots,i_m} \times_1 N^{\sigma_1}_{t^1_{i_1}} \times \cdots \times_m N^{\sigma_m}_{t^m_{i_m}}.$$

When $r=t^1_{i_1}+\cdots+t^m_{i_m}-(m-1)$, the second series vanish for $N^{t^j_{i_j}}_{t^j_{i_j}}=O,\ j=1,2,\ldots,m$. Hence Theorem 2(i) is proved.

(ii) The proof of Theorem 1(ii) shows that there exists a solution like Theorem 2(ii)(a)(b) in Eq. (3). Next we will prove that the solution of Eq. (3) is unique regardless of the choice of arbitrary parameters.

Suppose that $\mathscr{Y}_{i_1,i_2,\ldots,i_m}$ is a solution of Eq. (3).

Given $k_1,k_2,\ldots,k_m$, we let $l_1=t^1_{i_1}-(k_1-u_1),\ldots,l_{q-1}=t^{q-1}_{i_{q-1}}-(k_{q-1}-u_{q-1})$, $l_q=t^q_{i_q}-(r+1)-k_q$, $l_{q+1}=t^{q+1}_{i_{q+1}}-(k_{q+1}-u_{q+1}),\ldots,l_{m-1}=t^{m-1}_{i_{m-1}}-(k_{m-1}-u_{m-1})$, $l_m=t^m_{i_m}-[k_m-(r-u_1-\cdots-u_{q-1}-u_{q+1}-\cdots-u_{m-1})]$ with $u_j=\max\{r-k_{j+1}-\cdots-k_{q-1}-k_{q+1}-\cdots-k_m-u_1-\cdots-u_{q-1}-u_{q+1}-\cdots-u_{j-1},0\},\ldots,\min\{r-u_1-\cdots-u_{q-1}-u_{q+1}-\cdots-u_{j-1},k_j\},\ j=1,2,\ldots,q-1,q+1,\ldots,m-1$ and $r=0,1,\ldots,\min\{k_1+\cdots+k_{q-1}+k_{q+1}+\cdots+k_m,t^q_{i_q}-1-(k_q+1)\}$ in Eq. (5). Using the method in the proof of necessity of Theorem 3.2(ii) induces the formulas of the solution in Theorem 2(ii)(a)(b). Besides, the proof of necessity in Theorem 1(ii) indicates that components

$$\mathscr{Y}_{i_1,i_2,\ldots,i_m}(t^1_{i_1}-k_1,\ldots,t^{q-1}_{i_{q-1}}-k_{q-1},1,t^{q+1}_{i_{q+1}}-k_{q+1},\ldots,t^m_{i_m}-k_m),$$

the indexes in the parenthesis are the same as that of the variate of the system of non-homogeneous linear equations in Theorem 1(ii)(b), satisfy the system of non-homogeneous linear equations in Theorem 1(ii)(b).

There are apparently no other restrictions on the solution $\mathscr{Y}_{i_1,i_2,\ldots,i_m}$ since Eq. (3) has a solution like Theorem 2(ii)(a)(b). The proof is completed. □

Proof of Theorem 3

Proof.

(i) It is obvious for the uniqueness of the solution.

(ii) The number of equations in Eq. (5) is $t^1_{i_1}t^2_{i_2}\cdots t^m_{i_m}$.

(a) If $t^1_{i_1}+\cdots+t^{q-1}_{i_{q-1}}+t^{q+1}_{i_{q+1}}+\cdots+t^m_{i_m}-(m-1)<t^q_{i_q}$, then $k^1_{i_1}+\cdots+k^{q-1}_{i_{q-1}}+k^{q+1}_{i_{q+1}}+\cdots+k^m_{i_m}<t^q_{i_q},\ k_j=0,1,\ldots,t^j_{i_j}-1,\ j=1,\ldots,q-1,q+1,\ldots,m$. The number of equations in Theorem 1(ii)(a) is $t^1_{i_1}\cdots t^{q-1}_{i_{q-1}}t^{q+1}_{i_{q+1}}\cdots t^m_{i_m}$ and that in Theorem 3.2(ii)(b) is 0. Since Eq. (5) can change the components of $\mathscr{Q}_{i_1,i_2,\ldots,i_m}$ in the equations of Theorem 1(ii)(a) into linear combinations of the components of $\mathscr{Y}_{i_1,i_2,\ldots,i_m}$, Eq. (5) and the equations of Theorem 1(ii)(a) yield $t^1_{i_1}t^2_{i_2}\cdots t^m_{i_m}-t^1_{i_1}\cdots t^{q-1}_{i_{q-1}}t^{q+1}_{i_{q+1}}\cdots t^m_{i_m}$ equations with $t^1_{i_1}t^2_{i_2}\cdots t^m_{i_m}-1$ unknowns. Adding arbitrary element $\mathscr{Y}_{i_1,i_2,\ldots,i_m}(1,1,\ldots,1)$ obtains $t^1_{i_1}\cdots t^{q-1}_{i_{q-1}}t^{q+1}_{i_{q+1}}\cdots t^m_{i_m}$ arbitrary parameters in $\mathscr{Y}_{i_1,i_2,\ldots,i_m}$.

(b) If $t^1_{i_1}+\cdots+t^{q-1}_{i_{q-1}}+t^{q+1}_{i_{q+1}}+\cdots+t^m_{i_m}-(m-1)\geqslant t^q_{i_q}$, then both $k^1_{i_1}+\cdots+k^{q-1}_{i_{q-1}}+k^{q+1}_{i_{q+1}}+\cdots+k^m_{i_m}<t^q_{i_q}$ and $k^1_{i_1}+\cdots+k^{q-1}_{i_{q-1}}+k^{q+1}_{i_{q+1}}+\cdots+k^m_{i_m}\geqslant t^q_{i_q}$, $k_j=0,1,\ldots,t^j_{i_j}$, $j=1,\ldots,q-1,q+1,\ldots,m$, occur. The number of equations in Theorem 1(ii) is $t^1_{i_1}\cdots t^{q-1}_{i_{q-1}}t^{q+1}_{i_{q+1}}\cdots t^m_{i_m}$ with the coefficient matrix, its rank is r, of the system of non-homogeneous linear equations in Theorem 1(ii)(b). Utilizing an argument similar to the one used to the proof of Theorem 3(ii)(a) gets the dimension of the solution space.

Furthermore, due to "the inclusion-exclusion principle" [3], the number of equations in Theorem 1(ii)(b) is

$$t^1_{i_1}\cdots t^{q-1}_{i_{q-1}}t^{q+1}_{i_{q+1}}\cdots t^m_{i_m}-\sum_{n=0}^{t^q_{i_q}-1}\{C^{m-2}_{m+n-2}+\sum_{t=1}^{m-1}(-1)^t\sum_{\substack{j_1<\cdots<j_t\\ j_1,\ldots,j_t\in[m]\setminus\{q\}}}C^{m-2}_{m+n-2-\sum_{l=1}^t t^{j_l}_{i_{j_l}}}\},$$

where $C^{m-2}_{m+n-2-\sum_{l=1}^t t^{j_l}_{i_{j_l}}}$ is valid when $n\geqslant\sum_{l=1}^t t^{j_l}_{i_{j_l}}$.

The proof is completed. □

REFERENCES

[1] F. P. ALI BEIK, F. S. MOVAHED, S. AHMADI-ASL, *On the Krylov subspace methods based on tensor format for positive definite Sylvester tensor equations*, Numerical Linear Algebra with Applications, 2016, **23** (3): 444–466.

[2] Z. Z. BAI, G. H. GOLUB, M. K. NG, *Hermitian and skew-hermitian splitting methods for non-hermitian positive definite linear systems*, SIAM Journal on Matrix Analysis & Applications, 2004, **98** (1): 1–32.

[3] R. A. BRUALDI, *Introductory combinatorics*, fifth edition, China Machine Press, Beijing, 2009, 3.

[4] C. CHEN, L. LU, *A projection method and kronecker product preconditioner for solving Sylvester tensor equations*, Science China Mathematics, 2012, **55** (6): 12.

[5] L. GRASEDYCK, *Existence and computation of low kronecker-rank approximations for large linear systems of tensor product structure*, Computing, 2004, **72** (3–4): 247–265.

[6] T. G. KOLDA, B. W. BADER, *Tensor decompositions and applications*, SIAM Review, 2009, **51** (3): 455–500.

[7] D. KRESSNER, C. TOBLER, *Krylov subspace methods for linear systems with tensor product structure*, SIAM Journal on Matrix Analysis & Applications, 2009, **31** (4): 1688–1714.

[8] T. LI, C. H. FENG, X. F. ZHANG, *Two algorithms for solving generalized coupled Sylvester tensor equations*, Filomat, 2023, **37** (30): 10249–10264.

[9] B. W. LI, Y. S. SUN, D. W. ZHANG, *Chebyshev collocation spectral methods for coupled radiation and conduction in a concentric spherical participating medium*, Journal of Heat Transfer, 2009, **131** (6): 177–181.

[10] B. W. LI, S. TIAN, Y. S. SUN AND Z. M. HU, *Schur-decomposition for 3D matrix equations and its application in solving radiative discrete ordinates equations discretized by Chebyshev collocation spectral method*, Journal of Computational Physics, 2010, **229** (4): 1198–1212.

[11] T. LI, Q. W. WANG, X. F. DUAN, *Numerical algorithms for solving discrete Lyapunov tensor equation*, Journal of Computational and Applied Mathematics, 2019, **370** (3): 112676.

[12] S. RAGNARSSON, C. F. V. LOAN, *Block tensors and symmetric embeddings*, Linear Algebra and Its Applications, 2013, **438** (2): 853–874.

[13] S. RAGNARSSON, C. F. V. LOAN, *Block tensor unfoldings*, SIAM Journal on Matrix Analysis & Applications, 2011, **33** (1): 149–169.

[14] X. SHI, Y. WEI, S. LING, *Backward error and perturbation bounds for high order Sylvester tensor equation*, Linear and Multilinear Algebra, 2013, **61** (10): 1436–1446.
[15] X. F. ZHANG, W. DING, T. LI, *Tensor form of GPBiCG algorithm for solving the generalized Sylvester quaternion tensor equations*, Journal of the Franklin Institute, 2023, **360**: 5929–5946.
[16] X. F. ZHANG, T. LI, Y. G. OU, *Iterative solutions of generalized Sylvester quaternion tensor equations*, Linear and Multilinear Algebra 2024, **72** (7): 1259–1278.

(Received February 4, 2025)

Li Liang
School of Mathematics and Statistics
Zhejiang University of Science and Technology
Hangzhou, China
e-mail: HITLiangLi@163.com