Abstract

Multilinear systems play an important role in scientific calculations of practical problems. In this paper, we consider a tensor splitting method with a relaxed Anderson acceleration for solving the multilinear systems. The new method preserves nonnegativity for every iterative step and improves the existing ones. Furthermore, the convergence analysis of the proposed method is given. The new algorithm performs effectively for numerical experiments.

Key words. Tensor splitting, multilinear systems, Anderson acceleration, strong $\mathcal{M}$-tensor

MSC 15A48 15A69 65F10 65H10

1 Introduction

Let $\mathcal{A} = (a_{i_1i_2\ldots i_l})$ be an order $l$ dimension $n$ tensor and $b \in \mathbb{R}^n$. The multilinear systems were first given in [1]:

$$\mathcal{A}x^{l-1} = b,$$

where $x$ is the unknown vector to be solved. The $i$-th entry of the vector $\mathcal{A}x^{l-1}$ (see [2, 3]) is defined as follows:

$$(\mathcal{A}x^{l-1})_i = \sum_{i_2,\ldots,i_l=1}^n a_{ii_2\ldots i_l}x_{i_2}\ldots x_{i_l}, \quad i = 1, 2, \ldots, n,$$

where $x_i$ denotes the $i$-th component of a vector $x$. Since there are many applications such as tensor complementarity problems (e.g., see [1, 4]), numerical partial differential equations (e.g., see [6, 7]) and Bellman equation (e.g., see [6, 7]), the nonlinear equation (1.1) has
attracted many researchers’ attention, e.g., theoretical analysis of solutions (e.g., see [1, 8, 9]), numerical algorithms such as splitting type methods (e.g., see [10, 11]), Gradient-based iterations (e.g., see [1, 4, 5, 12–16]), dynamical approaches (e.g., see [7]) and so on. The majorization matrix of a tensor $\mathcal{A}$ is denoted by $M(\mathcal{A})_{i_{1}i_{2}} = b_{i_{1}i_{2}...i_{2}}$ with $i_{1}, i_{2} = 1, 2, ..., n$. For a vector $y = (y_{i})$, we define $y_{l}^{[l-1]} = (y_{1}^{[l-1]}, y_{2}^{[l-1]}, ..., y_{n}^{[l-1]})^{T}$. In [10], several tensor splitting methods (TSM) were given via setting $\mathcal{A} = \mathcal{E} - \mathcal{F}$:

$$x_{k} = (M(\mathcal{E})^{-1}\mathcal{F}x_{k-1} + M(\mathcal{E})^{-1}b)^{[l-1]}.$$  \hfill (1.3)

The iterative solution of which is a fixed point of the following iterative function

$$g_{\mathcal{E}}(x) = (M(\mathcal{E})^{-1}\mathcal{F}x^{l-1} + M(\mathcal{E})^{-1}b)^{[l-1]},$$  \hfill (1.4)

where $M(\mathcal{E})^{-1}\mathcal{F}$ is the iterative tensor of the splitting method (1.3). In order to accelerate the convergence of tensor splitting methods, a preconditioned splitting method was first proposed in [17], and then the preconditioned methods were further studied (e.g., see [18–20]). It is noted that tensor splitting type methods have a fine numerical performance in terms of the running time because of avoiding calculation of the Jacobian matrix (e.g., see [20]). Therefore, an improved tensor splitting method combing with the efficient acceleration technique is still an attractive topic.

Anderson acceleration is a powerful extrapolation technique to improve the fixed-point iteration for solving the linear or nonlinear systems (e.g., see [21–23].) Niu et al. [24] gave an Anderson-Richardson method for solving the equation (1.1). Though the given methods in [24] perform well, it is a pity that there is no theoretical analysis yet to ensure the nonnegativity of the iteration when the multilinear systems are solved. By this motivation, we propose a tensor splitting method with the relaxed Anderson acceleration and the nonnegativity is preserved in every new iterative step for solving the multilinear systems with the coefficient tensor being a strong $M$-tensor. Besides, the convergence analysis is given for preserving the nonnegativity. Furthermore, the proposed algorithm performs better than the existing preconditioned methods in numerical experiments.

The rest of this paper is given as follows. We introduce some definitions and lemmas in Section 2. Besides, the tensor splitting method with the relaxed Anderson acceleration is proposed in Section 3. Furthermore, the convergence analysis is considered in Section 4. We report numerical experiments to present the efficient performance of the proposed method in Section 5. The final section is the conclusion.

## 2 Preliminaries

An order $l$ and dimension $n$ tensor $\mathcal{A}$ with $n^l$ elements is defined by

$$\mathcal{A} = (a_{i_{1}...i_{l}}), a_{i_{1}...i_{l}} \in \mathbb{R}, i_{j} \in \langle n \rangle, j = 1, \cdots, l,$$
where $\mathbb{R}$ denotes the real field, $\langle n \rangle = \{1, \cdots, n\}$ and $n$ is a positive integer. We use $\mathbb{R}^{[l,n]}$ to denote the set of all order $l$ dimension $n$ real tensors. If $l = 1$, $\mathbb{R}^{[1,n]}$ reduces to $\mathbb{R}^n$, which is commonly used to denote the set of all dimension $n$ real vectors. More specifically, $\mathbb{R}_+^n (\mathbb{R}_{++}^n)$ is indicated as the set of all nonnegative (positive) vectors. Let 0 denote a null number, a null vector, a null matrix or a null tensor. Let $A, B \in \mathbb{R}^{[l,n]}$. The notation $A \geq B$ (>$B$) refers that every element of $A$ is no less than (strictly greater than) the corresponding one of $B$.

Qi [2] and Lim [25] independently gave the definition of tensor eigenvalues and eigenvectors, i.e., a complex number $\lambda$ and a nonzero complex vector $x$ are called as an eigenvalue and an eigenvector of $A \in \mathbb{R}^{[l,n]}$, respectively, if the following equation holds:

$$Ax^{l-1} = \lambda x^{l-1},$$

where $x^{l-1} = (x_{1}^{l-1}, \ldots, x_{n}^{l-1})^T$. Besides, $(\lambda, x)$ is called as an $H$-eigenpair if both $\lambda$ and $x$ are real. Furthermore, the spectral radius of $A$ is signified by $\rho(A) = \max\{|\lambda| | \lambda \in \sigma(A)\}$ with $\sigma(A)$ being the set of all eigenvalues of $A$. We call $I_l \in \mathbb{R}^{[l,n]}$ a unit tensor if its entries are given by

$$\delta_{i_1 \cdots i_l} = \begin{cases} 1, & i_1 = \cdots = i_l, \\ 0, & \text{else}. \end{cases}$$

$A \in \mathbb{R}^{[l,n]}$ is a $Z$-tensor if its off-diagonal entries are non-positive. Besides, $A$ is an $M$-tensor (see the works of [26, 27]) if there exists a nonnegative tensor $B$ and a positive real number $\eta \geq \rho(B)$ such that $A = \eta I_l - B$. Furthermore, if $\eta > \rho(B)$, $A$ is a strong $M$-tensor. Let $A \in \mathbb{R}^{[2,n]}$ and $B \in \mathbb{R}^{[k,n]}$. A product [28] $C = AB \in \mathbb{R}^{[k,n]}$ is introduced by

$$c_{j_2 \cdots i_k} = \sum_{j_2=1}^{n} a_{jj_2} b_{j_2 i_2 \cdots i_k}.$$

In [10], it is pointed out that if $A$ is a left-nonsingular tensor if and only if $A = M(A)I_l$ and $M(A)$ is nonsingular. If $A$ is a strong $M$-tensor, then $M(A)$ is a nonsingular $M$-matrix. Besides, the some concepts on tensor splittings are listed as follows.

**Definition 2.1.** [12] Let $A, E, F \in \mathbb{R}^{[l,n]}$.

1. $A = E - F$ is a splitting of $A$ if and only if $E$ is a left-nonsingular.

2. $A = E - F$ is a regular splitting of $A$ if and only if $E$ is left-nonsingular with $M(E)^{-1} \geq 0$, and $F \geq 0$.

3. $A = E - F$ is a weak regular splitting of $A$ if and only if $E$ is left-nonsingular with $M(E)^{-1} \geq 0$ and $M(E)^{-1}F \geq 0$.

4. $A = E - F$ is a convergent splitting if and only if $\rho(M(E)^{-1}F) < 1$. 

3
Remark 2.2. From the works of [10, 27], if $A$ is a $Z$-tensor, the following terms are equivalent:

(1) $A$ is a strong $M$-tensor.
(2) $A$ has a convergent (weak) regular splitting.
(3) All the (weak) regular splittings of $A$ are convergent.
(4) All the $H$-eigenvalue of $A$ are positive.
(5) All the real part of each eigenvalue of $A$ is positive.

For more spectral structure and property of a strong $M$-tensor, please refer to the works of [27]. The following lemmas on multilinear systems in the existing works will be used in Section 4.

Lemma 2.3. [1] Let $A$ be a strong $M$-tensor and $b \in \mathbb{R}_+^n$. Thus, the multilinear systems (1.1) have a unique positive solution.

Lemma 2.4. [10] Let $A$ be a strong $M$-tensor and $b \in \mathbb{R}_+^n$. Thus, there always exists a $x_0 \in \mathbb{R}_+^n$ with $0 < Ax_0^{-1} \leq b$.

3 Tensor splitting methods with Anderson acceleration

Let $f_E(x) = g_E(x) - x$. The tensor splitting methods with the relaxed Anderson acceleration is proposed for solving the multilinear systems (1.1) as follows.

Algorithm 1 Restarted tensor splitting methods with the relaxed Anderson acceleration (RTSMRAA)

1: Require: Given a positive vector $b$, a strong $M$-tensor $A$ with its (weak) regular splitting $A = E - F$, maximum $k_{\text{max}}$, a positive initial vector $z_0$ and $m \geq 1$.
2: Update $z_1 = g_E(z_0), \mu_1 = z_1, y_1 = z_1$ and $f_0 = z_1 - z_0$.
3: for $k=1,2,\ldots, k_{\text{max}}$ do
4: Set $m_k = \min\{m, k\}$.
5: Update
$$\mu_{k+1} = g_E(z_k),$$
$$f_k = \mu_{k+1} - z_k.$$ (3.1)
6: Set $Q_k = (f_{k-m_k}, \ldots, f_k)$.
7: Determine $\alpha^{(k)} = (\alpha_0^{(k)}, \ldots, \alpha_{m_k}^{(k)})^T$ that solves
$$\min_{\alpha=(\alpha_0, \ldots, \alpha_{m_k})^T} \|Q_k\alpha\| \text{ s.t. } \sum_{i=0}^{m_k} \alpha_i = 1.$$ (3.2)
8: Update
\[ y_{k+1} = \sum_{i=0}^{m_k} \alpha_i^{(k)} g \varepsilon(z_{k-m_k+i}). \]

9: if \( y_{k+1} \geq 0 \) and \( \sum_{i=0}^{m_k} |\alpha_i^{(k)}| \leq \kappa_\alpha \) then
10: Choose \( \theta_k \in [0,1] \) to update \( z_{k+1} = \theta_k y_{k+1} + (1 - \theta_k) \mu_{k+1}. \)
11: else
12: Update \( z_{k+1} = \mu_{k+1}. \)
13: end if
14: Update \( k \leftarrow k + 1. \)
15: end for
16: return \( z_k \)

Remark 3.1. The optimization problem (3.2) can be solved by the following equivalent the square problem [23]: Give \( \zeta_i^{(k)} = (\zeta_0^{(k)}, \zeta_1^{(k)}, ..., \zeta_{m_k-1}^{(k)})^T \) by solving
\[ \min_{\zeta_i^{(k)}} \|f_k - F_k \zeta_i^{(k)}\|, \quad (3.3) \]
where \( F_k = (\Delta f_{k-m_k}, ..., \Delta f_{k-1}) \) and \( \Delta f_i = f_{i+1} - f_i. \) Furthermore, Step 8 in the RTSMRAA is written as follows:
\[ y_{k+1} = \mu_{k+1} - G_k \zeta_i^{(k)}, \quad (3.4) \]
where \( G_k = (\Delta g_{k-m_k}, ..., \Delta g_{k-1}) \) and \( \Delta g_i = g \varepsilon (z_{i+1}) - g \varepsilon (z_i). \) The solution \( \zeta \) of the least square problem (3.3) and the solution \( \alpha \) of the optimization problem (3.2) are related by \( \alpha_0 = \zeta_0, \alpha_i = \zeta_i - \zeta_{i-1} \) for \( i = 1, ..., m_k - 1 \) and \( \alpha_{m_k} = 1 - \zeta_{m_k-1}. \)

4 Convergence analysis

In this section, we will consider the convergence analysis of Algorithm 1. Let \( U_\delta \equiv \{x| ||x - x^*|| \leq \delta, \delta > 0\}. \) We write \( a_{i_1i_2...i_l} \) as \( a_{i\pi}, \) where \( \pi = i_2 \cdots i_l. \) Notice that for any \( A = (a_{i_1i_2...i_l}) \in \mathbb{R}^{[l,n]}, \) it is easy to check that
\[ Ax^{l-1} = \overline{A} x^{l-1} \]
where \( \overline{A} = (\overline{a}_{i_1i_2...i_l}) \) is given by:
\[ \overline{a}_{i_1i_2...i_l} = \frac{1}{(l-1)!} \sum_{\pi \in G} a_{i\pi}, \]
and \( G \) denotes the set of all permutations in \((i_2 \cdots i_l). \) Then \( \overline{A} \) is symmetric on the last \( l - 1 \) indices (\( \overline{A} \) is called semi-symmetric [29]). For the sake of clarity, we always denote the semi-symmetric of a tensor by the notation \( \overline{\tau} \) (e.g., \( \overline{A} \)). Firstly, we give some lemmas as follows.
Lemma 4.1. Let $A$ be a strong $M$-tensor, $b \in \mathbb{R}^n_{++}$ and $A = E - F$ be a (weak) regular splitting. Then the function $g_E(x)$ has the following properties:

(1) $g_E(x) \in \mathbb{R}^n_{++}$ for all $x \in \mathbb{R}^n$. 

(2) $g_E(x)$ is monotone increasing in $\mathbb{R}^n_{++}$, i.e., for any $0 \leq x \leq y$, $g_E(x) \leq g_E(y)$. 

Proof: Since $A = E - F$ is a (weak) regular splitting, the second assertion is obvious. Because $b \in \mathbb{R}^n_{++}$, it is obtained that $M(E)^{-1}b \in \mathbb{R}^n_{++}$. Thus, the first affirmation is proved.

Proposition 4.2. Under the assumption of Lemma 4.1, for any initial vector $z_0 \in \mathbb{R}^n_{++}$ with $0 < Az_0^{-1} \leq b$, we get $z_k \in \mathbb{R}^n_{++}$ generated by the RTSMRAA if $\theta_k \in [0, 1)$ in Step 10 of the RTSMRAA for all $k \geq 1$.

Proof: Since $Az_0^{-1} \leq b$, it is true that $z_0 \leq z_1$ under the assumption of Lemma 4.1. By Lemma 4.1, we know $\mu_k \in \mathbb{R}^n_{++}$ for any $k \geq 1$ via the induction. Then the proof is completed.

Remark 4.3. If $\theta_k = 1$ in Step 10 of the RTSMRAA for some $k$, it is noted that the TSM is just improved by the restarted Anderson acceleration for the $k$-th iteration, i.e., $z_k = y_k$. In this case, we still set $y_{k+1} \geq 0$ rather than $y_{k+1} > 0$ in Step 9 of the RTSMRAA. That’s because, by Lemma 4.1, we can still get a vector $z_{k+1} \in \mathbb{R}^n_{++}$. If $y_{k+1}$ is not nonnegative for all $k$, the RTSMRAA reduces to the TSM. Therefore, the next step $z_{k+1}$ can be always positive by Lemma 4.1.

Lemma 4.4. Let $x^*$ be the solution of the multilinear systems (1.1) with $A$ being a strong $M$-tensor and $b \in \mathbb{R}^n_{++}$. For every (weak) regular splitting $A = E - F$, there exists a $U_\delta$ such that the mapping $g_E(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ has the following properties for a sufficiently small $\delta$:

(i) For any $x \in U_\delta$, $g_E(x) \in U_\delta$.

(ii) For any $x, y \in U_\delta$, there exists a real number $c \in [0, 1)$ such that

$$\|g_E(x) - g_E(y)\| \leq c\|x - y\|.$$ 

(iii) For any $x \in U_\delta$,

$$\nabla g_E(x) = (E g_E(x)^{l-2})^{-1}F x^{l-2}, \ \nabla f_E(x) = \nabla g_E(x) - I.$$ 

(iv) $\nabla g_E(x^*)$ is Lipschitz continuously, i.e., for any $x, y \in U_\delta$, there exists a real number $\gamma > 0$ such that

$$\|\nabla f_E(x) - \nabla f_E(y)\| \leq \gamma\|x - y\|.$$
Proof: Since $A$ is a strong $\mathcal{M}$-tensor and $b$ is positive, by Lemma 2.3, $x^*$ is the unique positive solution. By the proof of Theorem 5.7 in [10], we have that $\mathcal{E}g_{\varepsilon}(x^*)^{[l-2]} = \mathcal{E}(x^*)^{[l-2]}$ is nonsingular and $\nabla g_{\varepsilon}(x) = (\mathcal{E}g_{\varepsilon}(x))^{[l-2]}\mathcal{F}x^{l-2}$ satisfying $\rho(\nabla g_{\varepsilon}(x)) \leq c < 1$. Because $g_{\varepsilon}(x)$ and $\nabla g_{\varepsilon}(x)$ are continuous, there exists $U_\delta$ such that $\rho(\nabla g_{\varepsilon}(x)) \leq c < 1$ and $\mathcal{E}g_{\varepsilon}(x)^{[l-2]}$ is nonsingular for any $x \in U_\delta$. Thus, it is easy to check that the affirmation (ii) holds by differential mean value theorem and the term (iii) is given. Besides, for any $x \in U_\delta$, we have

$$||g_{\varepsilon}(x) - x^*|| = ||g_{\varepsilon}(x) - g_{\varepsilon}(x^*)|| \leq c||x - x^*|| < ||x - x^*|| \leq \delta.$$  

Then the desired result (i) is obtained. Furthermore, by the result 3.3.5 of [30] and the second G-derivative $\nabla^2 f_{\varepsilon}(x)$ of $\nabla f_{\varepsilon}(x)$ at $x$ is continuous, we have

$$||\nabla f_{\varepsilon}(x) - \nabla f_{\varepsilon}(y)|| \leq \max_{x,y \in \mathbb{R}_+^n} \max_{t \in [0,1]} (||\nabla^2 f_{\varepsilon}(x + t(y - x))||)||x - y||.$$ 

Since $\mathbb{R}_+^n$ is a closed set, the term (iv) is shown by $\gamma = \max_{x,y \in \mathbb{R}_+^n} \max_{t \in [0,1]} (||\nabla^2 f_{\varepsilon}(x + t(y - x))||)||x - y||$. 

**Proposition 4.5.** Under the assumption of Lemma 4.4, we have $(\mathbb{I} \setminus \{x^*\}) \cap \left(U_\delta \setminus \{x^*\}\right) \neq \emptyset$, where $\mathbb{I} = \{x | x > 0, 0 < Ax^{l-1} \leq b\}$ and $U_\delta$ is defined in Lemma 4.4.

**Proof:** By Lemma 2.3, $\mathbb{I} \setminus \{x^*\} \neq \emptyset$. Let $p(x) = Ax^{l-1}$ which is a continuous function from $\mathbb{R}^n \to \mathbb{R}^n$. Thus, there exists a real number $\delta' > 0$ such that $0 < Ax^{l-1} \leq b$ for any $\tilde{x} \in U_\delta' \setminus \{x^*\}$. Let $\delta'' = \min\{\delta', \delta\}$. It is easy to check that $\tilde{x} \in U_\delta'' \subseteq U_\delta$. Therefore, the proof is completed. 

**Lemma 4.6.** Under the assumption of Lemma 4.4, there exists $U_\delta \subseteq U_\delta$ with sufficiently small $\delta$ such that for any $x \in \mathbb{R}_+^{n+}$,

$$||f_{\varepsilon}(x) - \nabla f_{\varepsilon}(x^*)(x - x^*)|| \leq \frac{1}{2} ||x - x^*||^2 \quad (4.1)$$ 

and

$$(1 - c)||x - x^*|| \leq ||f_{\varepsilon}(x)|| \leq (1 + c)||x - x^*||. \quad (4.2)$$ 

**Proof:** The formula (4.1) is a direct result by 3.2.12 of [30] and (iii) of Lemma 4.4. The second formula can be shown by (ii) of Lemma 4.4 and

$$||f_{\varepsilon}(x)|| = ||f_{\varepsilon}(x) - f_{\varepsilon}(x^*)|| = ||g_{\varepsilon}(x) - g_{\varepsilon}(x^*) + (x - x^*)||.$$ 

Let $\epsilon_k = z_k - x^*$, $\hat{\epsilon}_k = y_k - z^*$ and $z^* = x^*$. Then we give the convergence analysis of the proposed algorithm as follows.
Theorem 4.7. Under the assumption of Lemma 4.4, for any initial vector \(z_0 \in \mathbb{II} \cap U_\delta\), there exists \(U_\hat{\delta} \subseteq \mathbb{II} \cap U_\delta\) with a small enough \(\hat{\delta}\) such that \(y_k, z_k \in U_\hat{\delta}(k \geq 1)\), and

\[
\|f_E(z_k)\| \leq \tau^k \|f_E(z_0)\|,
\]

where \(y_k\) and \(z_k\) are generated by the RTSMRAA, \(\forall \tau \in (c, 1)\) and \(\tau = \frac{1+c}{2}\).

Proof: We can always choose the \(z_0 \in U_\hat{\delta}\) with a sufficiently small \(\hat{\delta}\) such that \(\hat{\delta} < 2(1-c)/\gamma\),

\[
\frac{\kappa_\alpha \tau^2(2c + \gamma \hat{\delta})}{2(1-c)} \tau^{-m-1} \|f_E(z_0)\| \leq \frac{\kappa_\alpha \tau^2(2c + \gamma \hat{\delta})(1+c)}{2(1-c)} \tau^{-m} \|\epsilon_0\| \leq \hat{\delta}
\]

and

\[
\frac{c + \frac{\kappa_\alpha \gamma \hat{\delta}}{2(1-c)} \tau^{-m-1}}{1 - \frac{\gamma}{2(1-c)}} \leq 1.
\]

For any initial vector \(z_0 \in U_\hat{\delta} \subseteq \mathbb{II} \cap U_\delta\) with a sufficiently small \(\hat{\delta}\) such that (4.1) and (4.2) hold, by Lemma 4.4 and the proof of Proposition 4.5, we know that there exists \(U_\hat{\delta} \subseteq \mathbb{II} \cap U_\delta\) such that \(y_1, y_1, \mu_1, \mu_2 \in U_\hat{\delta}\) and the inequality (4.3) holds. Assume that \(y_k, z_k \in U_\hat{\delta}\) and \(\|f_E(z_k)\| \leq \tau^k \|f_E(z_0)\|\) for any \(0 \leq k \leq N\). Now we will prove \(z_{N+1} \in U_\hat{\delta}\) and the inequality \(\|f_E(z_{N+1})\| \leq \tau^{N+1} \|f_E(z_0)\|\) by the induction. By Lemma 4.6, we have

\[
f_E(z_k) = \nabla f_E(z^\ast) \epsilon_k + \Omega_k,
\]

and \(\|\Omega_k\| \leq \frac{c}{2} \|\epsilon_k\|^2\). Therefore, it is easy to check that

\[
g_E(z_k) = z^\ast + \nabla g_E(z^\ast) \epsilon_k + \Omega_k.
\]

Firstly, we will prove (4.3) and \(y_{N+1} \in U_\hat{\delta}\) from the following two cases.

Case I: If \(y_{N+1} \leq 0\) or \(\sum_{i=0}^{m_k} |a_i^{(k)}| > \kappa_\alpha\), by Lemma 4.4, Lemma 4.6 and the hypothesis, we imply \(z_{N+1} = \mu_{N+1} \in U_\hat{\delta}\) and

\[
\|f_E(z_{N+1})\| \leq \|f_E(z_N)\| \leq \|f_E(z^\ast)\| \leq c(1+c) \|\epsilon_N\| \leq \frac{c(1+c)}{1-c} \|f_E(z_N)\| \leq \frac{1+c}{1-c} \tau^{N+1} \|f_E(z_0)\| = \tau^{N+1} \|f_E(z_0)\|.
\]

This gives the inequality (4.3).
Case II: If $y_{N+1} \geq 0$ and $\sum_{i=0}^{m_N} |\alpha^{(k)}_i| \leq \kappa, \alpha$, note that the fact $\sum_{i=0}^{m_N} \alpha^{(N)}_i = 1$ combing with (4.7) gives

$$y_{N+1} = z^* + \sum_{i=0}^{m_N} \alpha^{(N)}_i \nabla g(z^*) \epsilon_{N-m_N+i} + \sum_{i=0}^{m_N} \alpha^{(N)}_i \Omega_{N-m_N+i}. \quad (4.8)$$

By Lemma 4.6, we can obtain

$$\| \sum_{i=0}^{m_N} \alpha^{(N)}_i \Omega_{N-m_N+i} \| \leq \sum_{i=0}^{m_N} |\alpha^{(N)}_i| \| \Omega_{N-m_N+i} \| \leq \gamma \sum_{i=0}^{m_N} |\alpha^{(N)}_i| \| \epsilon_{N-m_N+i} \|^2 / 2. \quad (4.9)$$

Since $m_N = \min\{m, N\}$, we get $N - m_N + i \geq N - m$. Thus, by Lemma 4.6 and the assumption, it is not difficult to get that

$$\| \epsilon_{N-m_N+i} \| \leq \frac{1}{1 - c} \| f_{\epsilon}(z_{N-m_N+i}) \| \leq \frac{\tau}{1 - c} \bar{\epsilon}^{N-m_N+i} \| f_{\epsilon}(z_0) \| \leq \frac{\tau}{1 - c} \bar{\epsilon}^{-m} \| f_{\epsilon}(z_0) \| \quad (4.10)$$

and

$$\| \epsilon_{N-m_N+i} \|^2 \leq \frac{1}{1 - c} \| \epsilon_{N-m_N+i} \| \| f_{\epsilon}(z_{N-m_N+i}) \|$$

$$\leq \frac{\delta \tau}{1 - c} \bar{\epsilon}^{N-m_N+i} \| f_{\epsilon}(z_0) \|$$

$$\leq \frac{\delta \tau}{1 - c} \bar{\epsilon}^{-m} \| f_{\epsilon}(z_0) \|, \quad (4.11)$$

together with $\sum_{i=0}^{m_N} |\alpha^{(N)}_i| \leq \kappa, \alpha$, the inequalities (4.9) and (4.10), we imply that

$$\| \sum_{i=0}^{m_N} \alpha^{(N)}_i \Omega_{N-m_N+i} \| \leq \frac{\kappa \alpha \delta \tau}{2(1 - c)} \bar{\epsilon}^{-m} \| f_{\epsilon}(z_0) \| \leq \frac{\kappa \alpha \delta \tau}{2(1 - c)} \bar{\epsilon}^{-m} \| f_{\epsilon}(z_0) \| \quad (4.12)$$

and

$$\| \sum_{i=0}^{m_N} \alpha^{(N)}_i \nabla g(z^*) \epsilon_{N-m_N+i} \| \leq \frac{\kappa \alpha c \delta \tau}{1 - c} \bar{\epsilon}^{-m} \| f_{\epsilon}(z_0) \|. \quad (4.13)$$

From (4.8), we obtain

$$\hat{e}_{N+1} = \sum_{i=0}^{m_N} \alpha^{(N)}_i \nabla g(z^*) \epsilon_{N-m_N+i} + \sum_{i=0}^{m_N} \alpha^{(N)}_i \Omega_{N-m_N+i},$$

combing with (4.12), (4.13) and the condition (4.4) yields

$$\| \hat{e}_{N+1} \| \leq \frac{\kappa \alpha \tau (2c + \gamma \delta)}{2(1 - c)} \bar{\epsilon}^{-m} \| f_{\epsilon}(z_0) \| \leq \delta, \quad (4.14)$$
and hence \( y_{N+1} \in U_\delta \). Because \( \mu_{N+1} \in U_\delta \) by Lemma 4.4, it is found that \( z_{N+1} \in U_\delta \). Therefore, by Lemma 4.6 we can give the Taylor’s formula of \( f_\varepsilon(z) \) at the true solution \( z^* \) as follows.

\[
f_\varepsilon(z_{N+1}) = (\nabla g_\varepsilon(z^*) - I)e_{N+1} + O_{N+1}, \tag{4.15}
\]

with \( \|O_{N+1}\| \leq \frac{2}{\varepsilon}\|e_{N+1}\|^2 \). This combing with (4.8) and the fact \( \nabla g_\varepsilon(z^*) (\nabla g_\varepsilon(z^*) - I) = (\nabla g_\varepsilon(z^*) - I)^2 \) yields

\[
f_\varepsilon(z_{N+1}) = (\nabla g_\varepsilon(z^*) - I) (\theta_k e_{N+1} + (1 - \theta_k)(g_\varepsilon(z_N) - z^*)) + O_{N+1}
= \theta_k \nabla g_\varepsilon(z^*) \sum_{i=0}^{m_N} \alpha_i^{(N)} (\nabla g_\varepsilon(z^*) - I)e_{N-m_N+i} + \theta_k (\nabla g_\varepsilon(z^*) - I) \sum_{i=0}^{m_N} \alpha_i^{(N)} O_{N-m_N+i} + (1 - \theta_k)(\nabla g_\varepsilon(z^*) - I)(g_\varepsilon(z_N) - z^*) + O_{N+1}. \tag{4.16}
\]

By (4.9)-(4.12), (4.15) the inductive assumption and the fact \( \| \sum_{i=0}^{m_N} \alpha_i^{(N)} f_\varepsilon(z_{N-m_N+i}) \| \leq \| f_\varepsilon(z_N) \| \), it is obtained that

\[
\| \nabla g_\varepsilon(z^*) \sum_{i=0}^{m_N} \alpha_i^{(N)} (\nabla g_\varepsilon(z^*) - I)e_{N-m_N+i} + (\nabla g_\varepsilon(z^*) - I) \sum_{i=0}^{m_N} \alpha_i^{(N)} O_{N-m_N+i} \|
= \| \nabla g_\varepsilon(z^*) \sum_{i=0}^{m_N} \alpha_i^{(N)} (f_\varepsilon(z_{N-m_N+i}) - O_{N-m_N+i}) + (\nabla g_\varepsilon(z^*) - I) \sum_{i=0}^{m_N} \alpha_i^{(N)} O_{N-m_N+i} \|
= \| \nabla g_\varepsilon(z^*) \sum_{i=0}^{m_N} \alpha_i^{(N)} f_\varepsilon(z_{N-m_N+i}) - \sum_{i=0}^{m_N} \alpha_i^{(N)} O_{N-m_N+i} \|
\leq \| f_\varepsilon(z_N) \| + \| \sum_{i=0}^{m_N} \alpha_i^{(N)} O_{N-m_N+i} \|
\leq \left( \frac{c}{\varepsilon} + \frac{\kappa_\alpha \gamma \delta}{2(1-c)} \varepsilon^{m-1} \right) \tau \varepsilon^{N+1} \| f_\varepsilon(z_0) \| \tag{4.17}
\]

and

\[
\| (\nabla g_\varepsilon(z^*) - I)(g_\varepsilon(z_N) - z^*) \| = \| (\nabla g_\varepsilon(z^*) - I)f_\varepsilon(z_N) + (\nabla g_\varepsilon(z^*) - I)e_N \|
\leq \| \nabla g_\varepsilon(z^*) f_\varepsilon(z_N) \| + \| O_{N} \|
\leq \left( \frac{c}{\varepsilon} + \frac{\gamma \delta}{2(1-c)} \varepsilon^{m-1} \right) \tau \varepsilon^{N+1} \| f_\varepsilon(z_0) \|. \tag{4.18}
\]

Note that \( \|O_{N+1}\| \leq \frac{\gamma \delta}{2(1-c)} \| f_\varepsilon(z_{N+1}) \| \). By (4.16)-(4.18), the hypothesis (4.5) and the fact
\(\kappa_{\alpha} \geq 1\), it is not difficult to get that
\[
||f_{E}(z_{N+1})|| \leq \theta_{k} \frac{\kappa_{\alpha}}{1 - \gamma_{\delta}} \frac{c_{c} - m - 1}{1 - \frac{\gamma_{\delta}}{2(1-c)}} \tau c^{N+1} \|f_{E}(z_{0})\| + (1 - \theta_{k}) \frac{\kappa_{\alpha}}{1 - \gamma_{\delta}} \frac{c_{c} - m - 1}{1 - \frac{\gamma_{\delta}}{2(1-c)}} \tau c^{N+1} \|f_{E}(z_{0})\|
\]
\[
\leq \tau c^{N+1} ||f_{E}(z_{0})||.
\]
\(\text{(4.19)}\)

The proof is completed.

From Lemma 4.6 and Theorem 4.7, we can get the following corollary immediately.

**Corollary 4.8.** Under the assumption of Lemma 4.4, for any initial vector \(z_{0} \in \mathbb{I} \cap \mathbb{U}_{\delta}\), there exists \(U_{\hat{\delta}} \subseteq \mathbb{I} \cap \mathbb{U}_{\delta}\) with small enough \(\hat{\delta}\) such that
\[
||\epsilon_{k}|| \leq \tau c^{k} ||\epsilon_{0}||,
\]
where \(\tau\) and \(c\) are given in Theorem 4.7.

5 Numerical examples

All the numerical experiments will be done in MATLAB R2018a with the configuration: Intel(R) Core(TM) i9-11900K CPU 3.50GHz and 64.00G RAM. For all numerical examples, we use ‘CPU(s)’, ‘IT’ and ‘RES’ to denote the running time, the iterative step and the norm of a residual vector (i.e., \(\text{RES} = ||A_{z}^{-1} - b||\)) of the tested algorithms, respectively. We set the maximum iterative number as 1000, \(\kappa_{\alpha} = 1000\) and the stopping criteria as \(||A_{z}^{-1} - b|| < \varepsilon\) with \(\varepsilon = 10^{-11}\).

Let \(A = E - F\), \(D = DL_{I}\) and \(L = LL_{I}\), where \(D\) and \(L\) are the diagonal part and the strictly lower triangle part of \(M(A)\), respectively. We give several illustrations with different algorithms listed in Table 1 where ‘-’ means that it is unavailable.

| Methods | Abbreviations |
|---------|---------------|
| Jacobian-Anderson type method | \(D\) Jacobian-RAA(m) |
| GS-Anderson type method | \(D - L\) GS-RAA(m) |
| SOR-Anderson type method | \(\frac{1}{2}(D - \omega L)\) SOR-RAA(m) |
| Preconditioned method in [12] | \(D - L\) P_{max} |
| Preconditioned SOR-type method in [20] | \(\frac{1}{2}(D - \omega L)\) P_{SOR} |
| Preconditioned SOR-type method in [17] | \(\frac{1}{2}(D - \omega L)\) P_{SOR} |
| Preconditioned AOR-type method in [21] | \(\frac{1}{2}(D - \omega L)\) PAOR |
| Accelerated modified dynamical system method in [32] | - AMDS-POWER |
| Newton’s method in [1] | - Newton |

**Remark 5.1.** If \(m = 0\), it is noted that the numerical results are given by the TSM.

We search the optimal parameter \(\theta\) from 0.1 to 1 in the interval of 0.1, the parameter \(\omega\) from 1 to 2 in the interval of 0.1 for all the SOR-type methods, the parameters \(\alpha\) and \(\beta\)
Example 5.2. (§1) Taking order-3 dimension \(n\) nonsingular \(M\)-tensors \(A = sI - B\), where \(B\) is generated randomly by MATLAB, and \(s = (1 + \varepsilon) \max_{i=1,2,\ldots,n} (Be^2)i\), \(\varepsilon > 0\), and \(\varepsilon = 1\).

Example 5.3. Let \(A \in \mathbb{R}^{[3,n]}\). Then entries \(a_{111} = a_{nnn} = 8\). For \(i = 2, 3, \ldots, n-1\), \(a_{iii} = 8\), \(a_{i+1ii} = a_{ii-1i} = a_{ii+1} = -1/3\).

Example 5.4. (§4) Let \(s = \frac{n^2}{2}\) and \(A = sI - B \in \mathbb{R}^{[3,n]}\) with \(b_{ijk} = |\sin(i + j + k)|\).

In numerical experiments, let \(\theta = \theta_k\) for all \(k\). We set the right-hand side \(b = e\), where \(e = (1, 1, \ldots, 1)^T\). Furthermore, we take the initial value \(z_0 = e\) and \(z_0 = \frac{1}{n}e\) for Examples 5.2-5.3 and 5.4, respectively.

### Table 2: Numerical comparison of the proposed algorithms with the variation of \(m\).

| \(m\) | \(\text{CPU(s)}\) | IT | RES | \(\text{CPU(s)}\) | IT | RES | \(\text{CPU(s)}\) | IT | RES |
|------|-------------------|----|-----|-------------------|----|-----|-------------------|----|-----|
| Example 5.2 | 0 | 0.144271 | 39 | 9.38E-12 | 0.1424690 | 39 | 8.93E-12 | 0.0929389 | 25 | 5.10E-12 |
| | 1 | 0.0357574 | 9 | 6.57E-16 | 0.0350616 | 9 | 6.35E-16 | 0.0340594 | 9 | 6.35E-16 |
| | 2 | 0.0317609 | 8 | 1.63E-12 | 0.0300341 | 8 | 1.99E-12 | 0.0300594 | 8 | 1.99E-12 |
| | 3 | 0.0414601 | 11 | 6.16E-13 | 0.0389104 | 10 | 1.91E-12 | 0.0385751 | 10 | 1.91E-12 |
| | 4 | 0.0403306 | 11 | 6.38E-12 | 0.0396004 | 11 | 2.07E-12 | 0.0396004 | 11 | 1.15E-14 |
| Example 5.3 | 0 | 0.0552742 | 14 | 2.36E-12 | 0.0449658 | 12 | 1.95E-12 | 0.0445933 | 12 | 1.42E-12 |
| | 1 | 0.0378777 | 10 | 7.77E-12 | 0.0332321 | 9 | 4.01E-12 | 0.0318594 | 9 | 2.69E-12 |
| | 2 | 0.0374533 | 10 | 9.40E-12 | 0.0298825 | 8 | 6.82E-12 | 0.0297657 | 8 | 8.11E-12 |
| | 3 | 0.0373813 | 10 | 2.07E-12 | 0.0296058 | 8 | 7.43E-12 | 0.029158 | 8 | 6.62E-12 |
| | 4 | 0.0386935 | 10 | 1.94E-12 | 0.0303437 | 8 | 7.76E-12 | 0.0295858 | 8 | 5.54E-12 |
| Example 5.4 | 0 | 0.1466787 | 38 | 7.90E-12 | 0.1415589 | 38 | 7.63E-12 | 0.0794291 | 22 | 6.27E-12 |
| | 1 | 0.0224974 | 6 | 1.78E-16 | 0.0232325 | 6 | 1.75E-16 | 0.0223109 | 6 | 1.75E-16 |
| | 2 | 0.0228135 | 6 | 7.50E-13 | 0.0232898 | 6 | 3.31E-13 | 0.022825 | 6 | 3.31E-13 |
| | 3 | 0.0223159 | 6 | 1.88E-12 | 0.0223231 | 6 | 1.95E-12 | 0.0228104 | 6 | 1.95E-12 |
| | 4 | 0.0226190 | 6 | 5.68E-12 | 0.0229646 | 6 | 4.58E-12 | 0.0225216 | 6 | 4.58E-12 |

### Table 3: The optimal parameters of the proposed algorithms

| \(m\) | \(\text{Jacobian-RAA}(m)\) | \(\text{GS-RAA}(m)\) | \(\text{SOR-RAA}(m)\) |
|------|-----------------|-----------------|-----------------|
| \(\theta\) | \(\theta\) | \((\theta, \omega)\) | \((\theta, \omega)\) |
| Example 5.2 | 0 | \(-\) | \((-1, 1)\) | 
| | 1 | 1 | 1 | \((1, 1)\) |
| | 2 | 1 | 1 | \((1, 1)\) |
| | 3 | 1 | 0.9 | \((0.9, 1)\) |
| | 4 | 0.9 | 0.9 | \((0.9, 1)\) |
| Example 5.3 | 0 | \(-\) | \((-1, 1)\) |
| | 1 | 0.7 | 0.5 | \((0.8, 1)\) |
| | 2 | 0.5 | 0.7 | \((0.9, 1)\) |
| | 3 | 0.8 | 0.6 | \((0.9, 1)\) |
| | 4 | 0.9 | 0.6 | \((0.8, 1)\) |
| Example 5.4 | 0 | \(-\) | \((-1, 1)\) |
| | 1 | 1 | 1 | \((1, 1)\) |
| | 2 | 1 | 1 | \((1, 1)\) |
| | 3 | 1 | 1 | \((1, 1)\) |
| | 4 | 1 | 1 | \((1, 1)\) |

It is easy to check that \(A\) in Examples 5.2-5.4 is a strong \(M\)-tensor and the corresponding multilinear systems (1.1) have a unique solution. For \(n = 200\), we report numerical results
in Table 2 and the optimal parameters are listed in Table 3 where $-1$ means that it is unavailable. It is seen that if $m = 2$ and $m = 3$, the SOR-Anderson type method performs best for Examples 5.2 and 5.3, respectively. The proposed algorithm is valid for a just small $m$ and increasing the value of $m$ does not improve the algorithm any better. This is a good numerical behavior because a small value of $m$ means that there is not much computation involved for solving the least square problem [3,3].

| Table 4: Numerical comparison with different algorithms. |
|---------------------------------------------------------|
| CPU | IT | RES | CPU | IT | RES | CPU | IT | RES | CPU | IT | RES |
|----------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Jacobian-RAA(2) | 0.001117 | 8 | 1.49E-15 | 0.89E-14 | 3 | 9.38E-15 | 1.61E-14 | 3 | 6.34E-15 | 1.71E-14 | 3 | 3.10E-15 | 1.71E-14 |
| GS-RAA(2) | 0.000117 | 8 | 5.42E-15 | 3.51E-14 | 3 | 2.13E-15 | 3.11E-14 | 3 | 1.02E-15 | 3.11E-14 | 3 | 2.20E-15 | 3.11E-14 |
| SOR-RAA(2) | 0.000092 | 8 | 5.26E-15 | 3.51E-14 | 3 | 2.13E-15 | 3.11E-14 | 3 | 1.02E-15 | 3.11E-14 | 3 | 2.20E-15 | 3.11E-14 |
| P-RAA | 0.001203 | 27 | 9.44E-15 | 3.51E-14 | 3 | 2.13E-15 | 3.11E-14 | 3 | 1.02E-15 | 3.11E-14 | 3 | 2.20E-15 | 3.11E-14 |
| Newton | 0.000145 | 9 | 6.90E-15 | 3.00E-13 | 10 | 1.52E-15 | 6.39E-13 | 11 | 3.52E-15 | 6.39E-13 | 12 | 6.51E-15 | 6.39E-13 |
| P-RAA | 0.002347 | 40 | 3.13E-15 | 3.00E-13 | 10 | 1.52E-15 | 6.39E-13 | 11 | 3.52E-15 | 6.39E-13 | 12 | 6.51E-15 | 6.39E-13 |
| P_SOR | 0.002368 | 20 | 1.41E-15 | 3.00E-13 | 10 | 1.52E-15 | 6.39E-13 | 11 | 3.52E-15 | 6.39E-13 | 12 | 6.51E-15 | 6.39E-13 |
| Final | 0.002259 | 50 | 1.05E-15 | 6.31E-13 | 10 | 1.52E-15 | 6.39E-13 | 11 | 3.52E-15 | 6.39E-13 | 12 | 6.51E-15 | 6.39E-13 |
| AMDS-POWER(2) | 0.017400 | 12 | 1.47E-14 | 3.00E-13 | 10 | 1.52E-15 | 6.39E-13 | 11 | 3.52E-15 | 6.39E-13 | 12 | 6.51E-15 | 6.39E-13 |

| Table 5: The optimal parameters of different algorithms. |
|---------------------------------------------------------|
| Jacobian-RAA(m) | GS-RAA(m) | SOR-RAA(m) | P_SOR(20) | P_SOR(17) | PAOR(31) | AMDS-POWER(2) |
| (m, θ) | (m, θ) | (m, θ, ω) | (β, ω) | (α, ω) | (γ, β, ω, r) | β |
|----------------------------------|-----|-----|-----|-----|-----|-----|
| Example 2 | (2, 1) | (2, 1) | (2, 1, 1) | (1.4, 1.4) | (0.7, 1.4) | (1.2, -0.3, 1.1, 0.4) | 0.3 |
| Example 3 | (3, 0.8) | (3, 0.6) | (3, 0.9, 1.1) | (0.9, 1.1) | (1.3, 1.1) | (1.2, -0.1, 1.1, 0.8) | 0.26 |
| Example 5 | (3, 1) | (3, 1) | (3, 1, 1) | (0.9, 1.5) | (0.6, 1.5) | (1.6, -0.3, 1.3, 0.3) | 0.3 |

For different methods, the numerical results are shown in Table 4 and the optimal parameters are listed in Table 5. Besides, Figure 1 is given to show the relationship between the norm of a residual vector and the running time. It is seen that the SOR-Anderson type method performs better in running times than other tested algorithms.
Figure 1: The relationship between the residual and the running time with different algorithms.
6 Conclusion

The main contribution of this paper is aimed at proposing the restarted nonnegativity preserving tensor splitting methods by the relaxed Anderson acceleration for solving the multilinear systems with a strong $\mathcal{M}$-tensor and the local convergence analysis is also given. From the numerical experiments, it is seen that the proposed methods with a suitable parameter perform better than the tensor splitting methods without acceleration. Besides, the increasing value of $m$ does not improve the given algorithm better, and the proposed algorithm is valid just for a small value of $m$. However, how to find an optimal parameter $\theta_k$ is still an open problem, which is one of our future works.

Declarations

Funding The first author was supported in part by National Natural Science Foundation of China (No. 12101136), the Guangdong Basic and Applied Basic Research Foundations (No. 2023A1515011633), the Project of Science and Technology of Guangzhou (No. 2024A04J2056), the Opening Project of Guangdong Province Key Laboratory of Computational Science at the Sun Yat sen University(No. 2021004), the Open Project of Key Laboratory, School of Mathematical Sciences, Chongqing Normal University (No. CSSXK-FKTQ202002). The third author was funded by the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJQN202100505), the Natural Science Foundation Project of Chongqing (Grant No. cstc2021jcyj-msxmX0195), and the Program of Chongqing Innovation Research Group Project in University (Grant No. CXQT19018).

Acknowledgments The authors would like to thank the referees for their helpful comments and suggestions and Prof. Lei Du for sharing their work [24] with us.

Conflict of interest The authors declare that they have no conflict of interest.

References

[1] W. Ding, Y. Wei, Solving multi-linear system with $\mathcal{M}$-tensors, Journal of Scientific Computing, 68(2)(2016): 689-715.

[2] L. Qi, Eigenvalues of a real supersymmetric tensor, Journal of Symbolic Computation, 40(2005): 1302-1324.

[3] L. Qi, Z. Luo, Tensor analysis: Spectral theory and special tensors, Society for Industrial and Applied Mathematics, 2017.

[4] Z. Xie, X. Jin, Y. Wei, Tensor methods for solving symmetric $\mathcal{M}$-tensor systems, Journal of Scientific Computing, 74(1)(2018): 412-425.

[5] D. Li, J. Xu, H. Guan, Newton’s Method for $\mathcal{M}$-Tensor Equations, Journal of Optimization Theory and Applications, 190(2)(2021): 628-649.
[6] P. Azimzadeh, E. Bayraktar, High order Bellman equations and weakly chained diagonally dominant tensors, SIAM Journal on Matrix Analysis and Applications, 40(1)(2019): 276-298.

[7] X. Wang, M. Che, Y. Wei, Neural networks based approach solving multi-linear systems with $M$-tensors, Neurocomputing, 351(2019): 33-42.

[8] D. Li, H.B. Guan, X.Z. Wang, Finding a nonnegative solution to an $M$-tensor equation, Pacific Journal of Optimization, 16(3)(2020): 419-440.

[9] L. Liu, X. Li, S. Liu, Further study on existence and uniqueness of positive solution for tensor equations, Applied Mathematics Letters, 124(2022): 107686.

[10] D. Liu, W. Li, S. W. Vong, The tensor splitting with application to solve multi-linear systems, Journal of Computational and Applied Mathematics, 330(2018): 75-94.

[11] D. Li, S. Xie, H. Xu, Splitting methods for tensor equations, Numerical Linear Algebra with Applications, 24(5)(2017): e2102.

[12] L. Han, A homotopy method for solving multilinear systems with $M$-tensors, Applied Mathematics Letters, 69(2017): 49-54.

[13] H. He, C. Ling, L. Qi, G. Zhou, A globally and quadratically convergent algorithm for solving multilinear systems with $M$-tensors, Journal of Scientific Computing, 76(3)(2018): 1718-1741.

[14] X. Bai, H. He, C. Ling, G. Zhou, A nonnegativity preserving algorithm for multilinear systems with nonsingular $M$-tensors, Numerical Algorithms, 87(3)(2021): 1301-1320.

[15] M. Liang, B. Zheng, R. Zhao, Alternating iterative methods for solving tensor equations with applications, Numerical Algorithms, 80(4)(2019): 1437-1465.

[16] C. Lv, C. Ma, A Levenberg-Marquardt method for solving semi-symmetric tensor equations, Journal of Computational and Applied Mathematics, 332(2018): 13-25.

[17] W. Li, D. Liu, S.W. Vong, Comparison results for splitting iterations for solving multi-linear systems, Applied Numerical Mathematics, 134(2018): 105-121.

[18] L. Cui, M. Li, Y. Song, Preconditioned tensor splitting iterations method for solving multi-linear systems, Applied Mathematics Letters, 96(2019): 89-94.

[19] Y. Zhang, Q. Liu, Z. Chen, Preconditioned Jacobi type method for solving multi-linear systems with $M$-tensors, Applied Mathematics Letters, 104(2020): 106287.

[20] D. Liu, W. Li, S. W. Vong, A new preconditioned SOR method for solving multi-linear systems with an $M$-tensor, Calcolo, 57(2)(2020): 1-18.
[21] F. Lai, W. Li, X. Peng, Y. Chen, Anderson accelerated fixed-point iteration for multi-linear PageRank, Numerical Linear Algebra with Applications, (2023): e2499.

[22] A. Toth, C. T. Kelley, Convergence analysis for Anderson acceleration, SIAM Journal on Numerical Analysis, 53(2)(2015): 805-819.

[23] H. F. Walker, P. Ni, Anderson acceleration for fixed-point iterations, SIAM Journal on Numerical Analysis, 49(4)(2011): 1715-1735.

[24] J. Niu, L. Du, T. Sogabe, S. Zhang, A tensor Alternating Anderson-Richardson method for solving multilinear systems with $\mathcal{M}$-tensors, arXiv preprint [arXiv:2401.07602] (2024).

[25] L.-H. Lim, Singular values and eigenvalues of tensors: A variational approach, in CAMSAP'05: Proceeding of the IEEE International Workshop on Computational Advances in Multi-sensor Adaptive Processing, 2005: 129-132.

[26] W. Ding, L. Qi, Y. Wei, $\mathcal{M}$-tensors and nonsingular $\mathcal{M}$-tensors, Linear Algebra and Its Applications, 439(2013): 3264-3278.

[27] L. Zhang, L. Qi, G. Zhou, $\mathcal{M}$-tensors and some applications, SIAM Journal on Matrix Analysis and Applications, 35(2014): 437-452.

[28] T. G. Kolda, B. W. Bader, Tensor decompositions and applications, SIAM Review, 51(2009): 455-500.

[29] Q. Ni, L. Qi, A quadratically convergent algorithm for finding the largest eigenvalue of a nonnegative homogeneous polynomial map, Journal of Global Optimization, 61(4)(2015): 627-641.

[30] J. M. Ortega, W. C. Rheinboldt, Iterative solution of nonlinear equations in several variables. Classics Appl. Math. 30, Society for Industrial and Applied Mathematics, Philadelphia, 2000.

[31] Y. Chen, C. Li, A new preconditioned AOR method for solving multi-linear systems, Linear and Multilinear Algebra, (2023): 1-18.

[32] X. Wang, C. Mo, M. Che, Y. Wei, Accelerated dynamical approaches for finding the unique positive solution of $\mathcal{KS}$-tensor equations, Numerical Algorithms, 88(4)(2021): 1787-1810.