Noname manuscript No.
(will be inserted by the editor)

A globally and quadratically convergent algorithm for solving multilinear systems with $\mathcal{M}$-tensors

Hongjin He · Chen Ling · Liqun Qi · Guanglu Zhou

Received: date / Accepted: date

Abstract We consider multilinear systems of equations whose coefficient tensors are $\mathcal{M}$-tensors. Multilinear systems of equations have many applications in engineering and scientific computing, such as data mining and numerical partial differential equations. In this paper, we show that solving multilinear systems with $\mathcal{M}$-tensors is equivalent to solving nonlinear systems of equations where the involving functions are P-functions. Based on this result, we propose a Newton-type method to solve multilinear systems with $\mathcal{M}$-tensors. For a multilinear system with a nonsingular $\mathcal{M}$-tensor and a positive right side vector, we prove that the sequence generated by the proposed method converges to the unique solution of the multilinear system and the convergence rate is quadratic. Numerical results are reported to show that the proposed method is promising.

Keywords Multilinear systems · $\mathcal{M}$-tensor · Newton's method · quadratic convergence.

H. He · C. Ling
Department of Mathematics, School of Science, Hangzhou Dianzi University, Hangzhou, 310018, China.
E-mail: hehjmath@hdu.edu.cn

C. Ling
E-mail: macling@hdu.edu.cn

L. Qi
Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.
E-mail: maqilq@polyu.edu.hk

G. Zhou (✉)
Department of Mathematics and Statistics, Curtin University, Perth, Western Australia, Australia.
E-mail: g.zhou@curtin.edu.au

Mathematics Subject Classification (2010) 15A18 · 90C30 · 90C33

1 Introduction

Let $\Re$ be the real field and $\Re^n_+$ be the nonnegative orthant in $\Re^n$. The interior of $\Re^n_+$, consisting of vectors with positive coordinates, will be denoted by $\Re^n_{++}$. An m-th order n-dimensional tensor $\mathcal{A}$ consists of n^m entries in $\Re$, and it is defined as

$$\mathcal{A} = (a_{i_1 i_2 \ldots i_m}), \quad a_{i_1 i_2 \ldots i_m} \in \Re, \quad 1 \le i_1, i_2, \ldots, i_m \le n. \tag{1.1}$$

Throughout this paper, we denote the set of all real tensors of order m and dimension n by $\mathcal{T}^{m,n}$. For a given tensor $\mathcal{A} \in \mathcal{T}^{m,n}$, we call it a nonnegative if $a_{i_1 i_2 \ldots i_m} \ge 0$ for all indices $i_1, \ldots, i_m$, and denote it by $\mathcal{A} \ge 0$ for simplicity. Various applications of tensors, nonnegative tensor in particular, can be found in the most recent monograph [16]. Following the definition introduced in [15], we say that $\mathcal{A}$ is called *symmetric* if its entries $a_{i_1 i_2 \ldots i_m}$ are invariant under any permutation of their indices $\{i_1, i_2, \ldots, i_m\}$. In particular, for every index $i \in [n] := \{1, 2, \cdots, n\}$, if an $(m-1)$-th order n-dimensional square tensor $\mathcal{A}_i := (a_{i i_2 \ldots i_m})_{1 \le i_2, \ldots, i_m \le n}$ is symmetric, then $\mathcal{A}$ is called a *semi-symmetric* tensor with respect to the indices $\{i_2, \ldots, i_m\}$.

For $\mathcal{A} \in \mathcal{T}^{m,n}$ and a column vector $x := (x_1, x_2, ..., x_n)^\top \in \Re^n$, we define an n-dimensional column vector:

$$\mathcal{A}x^{m-1} := \left(\sum_{i_2, \ldots, i_m = 1}^{n} a_{i\, i_2 \ldots i_m} x_{i_2} \cdots x_{i_m} \right)_{1 \le i \le n}, \tag{1.2}$$

and $\mathcal{A}x^{m-2}$ denotes an $n \times n$ matrix defined by

$$\mathcal{A}x^{m-2} := \left(\sum_{i_3, \ldots, i_m = 1}^{n} a_{ij i_3 \ldots i_m} x_{i_3} \cdots x_{i_m} \right)_{1 \le i,j \le n}. \tag{1.3}$$

The spectral radius of $\mathcal{A}$, defined as $\rho(\mathcal{A})$, is the maximum modulus of the eigenvalues of $\mathcal{A}$. Let $\mathcal{I}$ be the m-order n-dimensional unit tensor whose entries are

$$\mathcal{I}_{i_1 i_2 \ldots i_m} = \begin{cases} 1, & \text{if } i_1 = i_2 = \ldots = i_m, \\ 0, & \text{otherwise.} \end{cases} \tag{1.4}$$

A tensor $\mathcal{A} \in \mathcal{T}^{m,n}$ is called a $\mathcal{Z}$-tensor if all of its off-diagonal entries are nonpositive. Moreover, $\mathcal{A}$ is called an $\mathcal{M}$-tensor [5,20] if it can be written as $\mathcal{A} = s\mathcal{I} - \mathcal{B}$ with $s \ge \rho(\mathcal{B})$, and particularly, $\mathcal{A}$ is called a nonsingular $\mathcal{M}$-tensor if $s > \rho(\mathcal{B})$.

In this paper, we are concerned with a multilinear system which is defined as

$$\mathcal{A}x^{m-1} = b, \tag{1.5}$$

where $\mathcal{A} \in \mathcal{T}^{m,n}$, $b \in \Re^n$, and $x = (x_1, x_2, ..., x_n)^\top \in \Re^n$ is an unknown vector. In recent years, it has been well documented that multilinear systems arise in a number of applications such as data mining and numerical partial differential equations, e.g., see recent papers [3, 6, 11, 12, 19], to name just a few.

As shown in the seminal work [6], one specific application of (1.5) is the numerical *partial differential equation* (PDE) with Dirichlet's boundary condition, i.e.,

$$\begin{cases} u(x)^{m-2} \cdot \Delta u(x) = -\tilde{f}(x) \ \text{ in } \boldsymbol{D}, \\ u(x) = \tilde{g}(x) \ \text{ on } \partial \boldsymbol{D}, \end{cases} \qquad (m = 3, 4, \ldots), \tag{1.6}$$

where $\Delta = \sum_{j=1}^{p} \frac{\partial^2}{\partial x_j^2}$ and $\boldsymbol{D} = [0,1]^p$. When $\tilde{f}(\cdot)$ is a constant function, this PDE is a nonlinear Klein-Gordon equation (see [13, 21]). As stated by Ding and Wei in [6], the elliptic problem $\Delta u(x) = -\tilde{f}(x)$ with Dirichlet's boundary condition $u(x) = \tilde{g}(x)$ can be discretized into a linear system of equations whose coefficient matrix is a nonsingular M-matrix

$$L_h^{(p)} = \sum_{j=0}^{p-1} \underbrace{I \otimes \cdots \otimes I}_{j} \otimes L_h \otimes \underbrace{I \otimes \cdots \otimes I}_{p-j-1},$$

where $h = 1/(n-1)$ and

$$L_h = \frac{1}{h^2} \begin{pmatrix} 1 & & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & & 1 \end{pmatrix} \in \Re^{n \times n}.$$

Similarly, the operator $u \mapsto u^{\vartheta} \cdot \Delta u$ can also be discretized into an mth-order nonsingular $\mathcal{M}$-tensor

$$\mathcal{L}_h^{(p)} = \sum_{j=0}^{p-1} \underbrace{\mathcal{I} \otimes \cdots \otimes \mathcal{I}}_{j} \otimes \mathcal{L}_h \otimes \underbrace{\mathcal{I} \otimes \cdots \otimes \mathcal{I}}_{p-j-1},$$

which satisfies $(\mathcal{L}_h \boldsymbol{u}^{m-1})_i = u_i^{m-2} \cdot (L_h \boldsymbol{u})_i$ for $i = 1, 2, \cdots, n$, where $\mathcal{L}_h$ is an mth-order $\mathcal{M}$-tensor with

$$\begin{cases} (\mathcal{L}_h)_{11\ldots1} = (\mathcal{L}_h)_{nn\ldots n} = 1/h^2, \\ (\mathcal{L}_h)_{ii\ldots i} = 2/h^2, \ \ i = 2, \ldots, n-1, \\ (\mathcal{L}_h)_{i(i-1)i\ldots i} = (\mathcal{L}_h)_{ii(i-1)\ldots i} = (\mathcal{L}_h)_{iii\ldots(i-1)} = 1/h^2(m-1), \ i = 2, \ldots, n-1, \\ (\mathcal{L}_h)_{i(i+1)i\ldots i} = (\mathcal{L}_h)_{ii(i+1)\ldots i} = (\mathcal{L}_h)_{iii\ldots(i+1)} = -1/h^2(m-1), \ i = 2, \ldots, n-1. \end{cases}$$

Accordingly, the PDE in (1.6) is discretized into an $\mathcal{M}$-equation $\mathcal{L}_h^{(p)}\boldsymbol{u}^{m-1} = \tilde{\boldsymbol{f}}$, which can be regarded as a higher order generalization of the one discussed in [10, 18]. Here, we also refer the reader to [11] for another specific real-life example on a particle's movement under the gravitation.

Taking a close look at the problem under consideration, it is clear that (1.5) can be regarded as a special nonlinear equation, which thus can be solved directly by existing solvers designed for general nonlinear equations. However, such a treatment way to (1.5) often ignores the multilinearity of tensors, which also encourages us to design structure-exploiting algorithms for finding solutions to (1.5). To our knowledge, the development of algorithmic design for (1.5) is still in its infancy. Recently, it has been shown by Ding and Wei [6] that the multilinear system (1.5) has a unique positive solution (i.e., all entries of the solution are positive) if $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor and b is a positive vector (i.e., $b \in \Re^n_{++}$). In [6,19], the authors proposed a Newton's method and some tensor methods for finding the solution of (1.5) when $\mathcal{A}$ is a symmetric tensor. However, when $\mathcal{A}$ is not symmetric, it is unknown whether or not their methods still work. Most recently, for any nonsingular $\mathcal{M}$-tensor $\mathcal{A}$, Han [9] proposed a homotopy method for finding the unique positive solution of (1.5) and proved its convergence to the desired solution. As we know, it is unclear whether or not the homotopy method has a superlinear convergence property. In this paper, we consider the multilinear system (1.5) without the symmetry property on $\mathcal{A}$, and propose a globally convergent Newton-type method, which has locally quadratic convergence rate. A series of numerical results show that our algorithm is stable and fast for random synthetic examples.

The rest of this paper is organized as follows. In Section 2, we first give an equivalent formulation for the multilinear system (1.5). In particular, we show that solving (1.5) is equivalent to solving a nonlinear system $W(y) = 0$ (see (2.3)), where the function W is a P-function on $\Re^n_{++}$, in addition to proving that the Jacobian of W at any $y \in \Re^n_{++}$ is nonsingular. In Section 3, we propose a Newton-type method to solve the multilinear system (1.5) and prove that the proposed method converges quadratically and globally if $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor and b is a positive vector. In Section 4, we report our numerical results to show the efficiency of the proposed method. Finally, we conclude the paper with some remarks in Section 5.

Notation. We conclude this section with some notation and terminology. For a continuously differentiable function $F : \Re^n \to \Re^n$, we denote the Jacobian of F at $x \in \Re^n$ by $F'(x)$. For $x \in \Re^n$, the 2-norm is denoted by $\|x\|$, and $\mathrm{diag}(x)$ denotes the $n \times n$ diagonal matrix generated by x. For $x, y \in \Re^n$, we define $x \odot y := (x_1y_1, x_2y_2, \ldots, x_ny_n)^\top$. Moreover, for given a scalar t and a vector $y \in \Re^n$, we use (t, y) to represent the $(n+1)$-dimensional column vector $(t, y^\top)^\top$ for notational simplicity.

2 An Equivalent Formulation of (1.5)

In this section, we give an equivalent formulation for the multilinear system (1.5). In particular, we reformulate (1.5) as a nonlinear system of equations where the involving function is a P-function on $\Re^n_{++}$. This is a vital step in the development of a quadratically convergent method for (1.5).

An $n \times n$ matrix $A = (a_{ij})$ is called nonnegative (or positive), denoted by $A \geq 0$ (or $A > 0$), if $a_{ij} \geq 0$ (or $a_{ij} > 0$) for all i and j. A is called a Z-matrix if all its off-diagonal entries are nonpositive. Any Z-matrix can be written as $sI - B$ with $B \geq 0$; it is called a nonsingular M-matrix if $s > \rho(B)$, and a singular M-matrix if $s = \rho(B)$, where $\rho(B)$ is the spectral radius of B. For M-matrices, from [2], we have the following theorem.

Theorem 2.1 *For a Z-matrix $A \in \Re^{n\times n}$, the following are equivalent:*

(i) *A is a nonsingular M-matrix .*
(ii) *$Av \in \Re^n_{++}$ for some vector $v \in \Re^n_{++}$.*
(iii) *All the principal minors of A are positive.*

An $n \times n$ matrix $A = (a_{ij})$ is called a P-matrix if all the principal minors of A are positive. A is called a P_0-matrix if all the principal minors of A are nonnegative. Clearly, a P-matrix is a P_0-matrix, and by Theorem 2.1, we have that a nonsingular M-matrix is a P-matrix.

Definition 2.1 ([8]) A function $F : K \subseteq \Re^n \to \Re^n$ is called a

(i) P_0-function on K if for all $x, y \in \Re^n$ with $x \neq y$, there is an index $i_0 = i_0(x, y)$ with

$$x_{i_0} \neq y_{i_0} \quad \text{and} \quad (x_{i_0} - y_{i_0})[F_{i_0}(x) - F_{i_0}(y)] \geq 0;$$

(ii) P-function on K if for all $x, y \in \Re^n$ with $x \neq y$, it holds that

$$\max_i (x_i - y_i)[F_i(x) - F_i(y)] > 0.$$

Theorem 2.2 ([8]) *Let $F : \Omega \supset K \subseteq \Re^n \to \Re^n$ be continuously differentiable on the open set Ω containing the set K.*

(i) *If $F'(x)$ is a P-matrix for all $x \in K$, then F is a P-function on K.*
(ii) *If $F'(x)$ is a P_0-matrix for all $x \in K$, then F is a P_0-function on K.*

For a given $y \in \Re^n_{++}$, let $y^{[\frac{1}{m}]} = (y_1^{\frac{1}{m}}, y_2^{\frac{1}{m}}, ..., y_n^{\frac{1}{m}})^\top$ and

$$D = \text{diag}\left(y^{[\frac{1}{m}-1]}\right), \tag{2.1}$$

where m is the order of $\mathcal{A}$. Then, we have

$$D^{-1} = \text{diag}\left(y^{[1-\frac{1}{m}]}\right).$$

Define $W : \Re^n_{++} \to \Re^n$ by

$$W(y) := D\mathcal{A}\left(y^{\left[\frac{1}{m}\right]}\right)^{m-1} - Db. \tag{2.2}$$

We consider the following nonlinear system of equations:

$$W(y) = 0. \tag{2.3}$$

Clearly, we have the following result.

Proposition 2.1 *If $y^* \in \Re^n_{++}$ is a solution of (2.3) then $x^* = (y^*)^{\left[\frac{1}{m}\right]}$ is a solution of (1.5). Conversely, if x^* is a positive solution of (1.5) then $y^* = (x^*)^{[m]}$ is a positive solution of (2.3).*

This proposition shows that (2.3) has a unique positive solution when $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor and b is a positive vector.

In the following, we will show that for any $y \in \Re^n_{++}$, the Jacobian of the function W defined in (2.2) is a nonsingular M-matrix.

Lemma 2.1 *Suppose that $\mathcal{A} \in \mathcal{T}^{m,n}$ is a $\mathcal{Z}$-tensor and $b \in \Re^n_{++}$. Then, for any $y \in \Re^n_{++}$, $W'(y)$ is a nonsingular M-matrix.*

Proof Since $\mathcal{A}$ is a $\mathcal{Z}$-tensor, there exist a scalar s and a nonnegative tensor $\mathcal{B} \in \mathcal{T}^{m,n}$ such that $\mathcal{A} = s\mathcal{I} - \mathcal{B}$. As a consequence, $W(y)$ defined by (2.2) can be written as

$$W(y) = s\mathbf{1} - \mathrm{diag}\left(y^{\left[\frac{1-m}{m}\right]}\right)\mathcal{B}\left(y^{\left[\frac{1}{m}\right]}\right)^{m-1} - \mathrm{diag}\left(y^{\left[\frac{1-m}{m}\right]}\right)b, \tag{2.4}$$

where $\mathbf{1} := (1, 1, \ldots, 1)^\top \in \Re^n$. Note that, for a given $\mathcal{B} \in \mathcal{T}^{m,n}$, there always exists a semi-symmetric tensor $\bar{\mathcal{B}} \in \mathcal{T}^{m,n}$ such that

$$\mathcal{B}x^{m-1} = \bar{\mathcal{B}}x^{m-1}, \qquad x \in \Re^n.$$

Let $F(x) := \bar{\mathcal{B}}x^{m-1}$. Then, it follows from the semi-symmetry of $\bar{\mathcal{B}}$ that

$$F'(x) = (m-1)\bar{\mathcal{B}}x^{m-2}.$$

Consequently, by simple computation on (2.4), we have

$$\begin{aligned}
W'(y) = &-\left(\frac{1-m}{m}\right)\mathrm{diag}\left(y^{\left[\frac{1}{m}-2\right]}\right)\mathrm{diag}\left(\bar{\mathcal{B}}\left(y^{\left[\frac{1}{m}\right]}\right)^{m-1}\right)\\
&-\mathrm{diag}\left(y^{\left[\frac{1}{m}-1\right]}\right)(m-1)\bar{\mathcal{B}}\left(y^{\left[\frac{1}{m}\right]}\right)^{m-2}\frac{1}{m}\mathrm{diag}\left(y^{\left[\frac{1}{m}-1\right]}\right)\\
&-\left(\frac{1-m}{m}\right)\mathrm{diag}\left(y^{\left[\frac{1}{m}-2\right]}\right)\mathrm{diag}(b)\\
= &\left(\frac{m-1}{m}\right)\mathrm{diag}\left(y^{\left[\frac{1}{m}-2\right]}\right)\mathrm{diag}\left(\bar{\mathcal{B}}\left(y^{\left[\frac{1}{m}\right]}\right)^{m-1}\right)
\end{aligned}$$

$$-\left(\frac{m-1}{m}\right)\operatorname{diag}\left(y^{[\frac{1}{m}-1]}\right)\bar{\mathcal{B}}\left(y^{[\frac{1}{m}]}\right)^{m-2}\operatorname{diag}\left(y^{[\frac{1}{m}-1]}\right)$$
$$+\left(\frac{m-1}{m}\right)\operatorname{diag}\left(y^{[\frac{1}{m}-2]}\right)\operatorname{diag}(b), \tag{2.5}$$

which clearly shows that $W'(y)$ is a Z-matrix. Next, we show $W'(y)$ is a nonsingular M-matrix for any $y \in \Re^n_{++}$. Invoking the definition of D given in (2.1), we define a matrix M of the form

$$\begin{aligned} M &:= \frac{m}{m-1}D^{-1}W'(y)D^{-1} \\ &= \operatorname{diag}\left(y^{[-\frac{1}{m}]}\right)\operatorname{diag}\left(\bar{\mathcal{B}}\left(y^{[\frac{1}{m}]}\right)^{m-1}\right) - \bar{\mathcal{B}}\left(y^{[\frac{1}{m}]}\right)^{m-2} + \operatorname{diag}\left(y^{[-\frac{1}{m}]}\right)\operatorname{diag}(b) \\ &= \operatorname{diag}\left(y^{[-\frac{1}{m}]}\odot\bar{\mathcal{B}}\left(y^{[\frac{1}{m}]}\right)^{m-1}\right) - \bar{\mathcal{B}}\left(y^{[\frac{1}{m}]}\right)^{m-2} + \operatorname{diag}\left(y^{[-\frac{1}{m}]}\odot b\right). \end{aligned} \tag{2.6}$$

It is clear from the nonnegativeness of $\bar{\mathcal{B}}$ that M is a Z-matrix. Moreover,

$$My^{[\frac{1}{m}]} = \bar{\mathcal{B}}\left(y^{[\frac{1}{m}]}\right)^{m-1} - \bar{\mathcal{B}}\left(y^{[\frac{1}{m}]}\right)^{m-2}y^{[\frac{1}{m}]} + b = 0 + b = b \in \Re^n_{++}, \tag{2.7}$$

which, together with the equivalence of Item (i) and (ii) of Theorem 2.1, means that M is a nonsingular M-matrix. Additionally, from the definition of M in (2.6), we can see that $W'(y) = \hat{c}DMD$ with $\hat{c} := \frac{m-1}{m}$. Consequently, for any $y \in \Re^n_{++}$, it follows from (2.7) and $\hat{c} > 0$ that

$$W'(y)y = \hat{c}DMDy = \hat{c}DMy^{[\frac{1}{m}]} = \hat{c}Db = \hat{c}y^{[\frac{1}{m}-1]} \odot b \in \Re^n_{++}.$$

By using Theorem 2.1 again, we conclude that $W'(y)$ is a nonsingular M-matrix for any $y \in \Re^n_{++}$. □

Lemma 2.2 *Suppose that $\mathcal{A} \in \mathcal{T}^{m,n}$ is a $\mathcal{Z}$-tensor and $b \in \Re^n_{++}$. Then, $W(y)$ defined in (2.2) is P-function in $\Re^n_{++}$.*

Proof First, by the definition of M in (2.6), M is a nonsingular M-matrix, which means that $Mv \in \Re^n_{++}$ for all $v \in \Re^n_{++}$. Then, all the principle minors of M are positive. Hence, M is a P-matrix, which means $W'(y)$ is a P-matrix for any $y \in \Re^n_{++}$. So, $W(y)$ is P-function in $\Re^n_{++}$. □

Based on Lemmas 2.1 and 2.2, we may use the classic Newton method to solve (2.3). However, for some $c > 0$, the level set

$$L_W(c) := \{y \in \Re^n_{++} : \|W(y)\| \le c\}$$

may be unbounded even for multilinear systems with nonsingular $\mathcal{M}$-tensors and positive right side vectors. For example, let $\mathcal{A}$ be a 3-order 2-dimensional tensor with $a_{111} = a_{222} = 1$ and other elements being zero, and $b = (1,1)^\top$. Then,

$W(y) = (1 - y_1^{-2/3}, 1 - y_2^{-2/3})^\top$. Clearly, for $c = 2$, $L_W(c)$ is unbounded. This means when we use the Newton method to solve (2.3), the generated sequence may be unbounded. In order to overcome this, in the next section, we will propose a regularized Newton method for (2.3). For a multilinear system with a nonsingular $\mathcal{M}$-tensor and a positive right side vector, we will prove that the sequence generated by the proposed algorithm is bounded and converges to the unique solution of the multilinear system. Moreover, the convergence rate is quadratic.

3 A quadratically convergent algorithm

In this section, we will first present a Newton-type method for (1.5). Then, we prove that our new algorithm is globally and quadratically convergent to the unique solution of (1.5) when $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor and b is a positive vector.

Define $H : \Re^{n+1} \to \Re^{n+1}$ by

$$H(t, y) := \begin{pmatrix} t \\ G(t, y) \end{pmatrix}, \tag{3.1}$$

where

$$G(t, y) = W(y) + ty. \tag{3.2}$$

Here, the function G is a regularized form of the function W and H is an augmented function of G. Clearly, $H(t^*, y^*) = 0$ if and only if $t^* = 0$ and $W(y^*) = 0$. Hence, finding a positive solution of (2.3) is equivalent to finding a solution of $H(t, y) = 0$.

For any t and $y \in \Re^n_{++}$, by simple computation, we have

$$H'(t, y) := \begin{pmatrix} 1 & 0 \\ G'_t(t, y) & G'_y(t, y) \end{pmatrix}, \tag{3.3}$$

where

$$G'_t(t, y) = y \quad \text{and} \quad G'_y(t, y) = W'(y) + tI_n. \tag{3.4}$$

Here, I_n is the $n \times n$ identity matrix.

Proposition 3.1 *For any $t \geq 0$ and $y \in \Re^n_{++}$, if $\mathcal{A} \in \mathcal{T}^{m,n}$ is a $\mathcal{Z}$-tensor and $b \in \Re^n_{++}$ then $H'(t, y)$ defined by (3.3) is nonsingular.*

Proof For $y \in \Re^n_{++}$ it follows from Lemma 2.1 that $W'(y)$ is nonsingular, which, together with $t \geq 0$ and (3.4), immediately shows that $G'_y(t, y)$ is nonsingular. So, it follows from (3.3) that $H'(t, y)$ is nonsingular. □

Choose $\bar{t} \in \Re_{++}$ and $\gamma \in (0, 1)$ such that $\gamma\bar{t} < 0.5$. Define the merit function $\psi : \Re^{n+1} \to \Re_{+}$ by

$$\psi(t, y) := \|H(t, y)\|^2$$

and define $\beta : \Re^{n+1} \to \Re_{+}$ by

$$\beta(t, y) := \gamma \min\{1, \psi(t, y)\}. \tag{3.5}$$

With the above preparations, we now present a Newton-type method to solve the system $H(t, y) = 0$, where H is defined in (3.1), which is stated as Algorithm 1:

Algorithm 1 (Quadratically Convergent Algorithm (QCA) for (1.5)).

1: Choose constants $\delta \in (0, 1)$ and $\sigma \in (0, \frac{1}{2})$. Let $t^0 := \bar{t} > 0$ and $y^0 \in \Re^n_{++}$ be starting points.
2: **while** $\|H(t^k, y^k)\| \neq 0$ **do**
3: Let $\beta_k := \beta(t^k, y^k)$.
4: Let $\Delta t^k = -t^k + \beta_k \bar{t}$. Compute Δy^k by solving the following linear system of equations:

$$G'_y(t^k, y^k)\Delta y^k = -G(t^k, y^k) - G'_t(t^k, y^k)\Delta t^k. \tag{3.6}$$

5: Find l_k being the smallest nonnegative integer l satisfying $y^k + \delta^{l_k}\Delta y^k \in \Re^n_{++}$ and

$$\psi\left(t^k + \delta^{l_k}\Delta t^k,\ y^k + \delta^{l_k}\Delta y^k\right) \leq \left[1 - 2\sigma\,(\,1 - \gamma\bar{t}\,)\,\delta^{l_k}\right]\psi(t^k, y^k). \tag{3.7}$$

6: Let $t^{k+1} := t^k + \delta^{l_k}\Delta t^k$ and $y^{k+1} := y^k + \delta^{l_k}\Delta y^k$.
7: **end while**

Remark 3.1 Algorithm 1 can be regarded as a regularized version of the Newton methods proposed in [17]. Regularized Newton methods for nonlinear complementarity problems and variational inequality problems have been studied in [7,14]. As we observed, solving the linear system of equations (3.6) (i.e., Step 4) dominates the main computational task of Algorithm 1. However, we will show that $G'_y(t^k, y^k)$ is always nonsingular, which means that (3.6) is well-defined. Therefore, we can employ many efficient solvers to solve it (see Section 4).

In what follows, we show that the method (Algorithm 1) has global and quadratic convergence properties when $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor and b is a positive vector. Some results are modified from the corresponding results in [7,14, 17].

Lemma 3.1 *Suppose that for some $(\tilde{t}, \tilde{y}) \in \Re_{++} \times \Re^n_{++}$, $H'(\tilde{t}, \tilde{y})$ is nonsingular. Then, there exist a closed neighbourhood $\mathcal{N}(\tilde{t}, \tilde{y})$ of $(\tilde{t}, \tilde{y})$ in $\Re_{++} \times \Re^n_{++}$ and a positive* *number* $\bar{\alpha} \in (0,1]$ *such that for any* $(t,y) \in \mathcal{N}(\tilde{t},\tilde{y})$ *and all* $\alpha \in [0,\bar{\alpha}]$, *it holds that* $H'(t,y)$ *is invertible and*

$$\psi(t+\alpha\Delta t, y+\alpha\Delta y) \le [1-2\sigma(1-\gamma\bar{t}\ )\alpha]\psi(t,y), \tag{3.8}$$

where $\Delta t = -t+\beta(t,y)\bar{t}$ *and* Δy *is a solution of the following linear system of equations:*

$$G'_y(t,y)\Delta y = -G(t,y) - G'_t(t,y)\Delta t. \tag{3.9}$$

Proof Since $H'(\tilde{t},\tilde{y})$ is invertible and $(\tilde{t},\tilde{y}) \in \Re_{++}\times\Re^n_{++}$, there exists a closed neighbourhood $\mathcal{N}(\tilde{t},\tilde{y})$ of $(\tilde{t},\tilde{y})$ in $\Re_{++}\times\Re^n_{++}$ such that for any $(t,y)\in\mathcal{N}(\tilde{t},\tilde{y})$, we have $H'(t,y)$ defined by (3.3) is invertible. For any $(t,y)\in\mathcal{N}(\tilde{t},\tilde{y})$, let $\Delta t = -t+\beta(t,y)\bar{t}$ and let Δy be a solution of the linear system of equations (3.9). Let $\Delta z := (\Delta t, \Delta y) \in \Re^{n+1}$ and $r := (\beta(t,y)\bar{t}, \mathbf{0}) \in \Re^{n+1}$ with $\mathbf{0} = (0,\ldots,0)^\top \in \Re^n$. Then, it follows from Step 4 of Algorithm 1 that

$$H(t,y) + H'(t,y)\Delta z = r, \quad \|r\| = \beta(t,y)\bar{t}. \tag{3.10}$$

For any $\alpha\in[0,1]$, define

$$g_{(t,y)}(\alpha) = H(t+\alpha\Delta t, y+\alpha\Delta y) - H(t,y) - \alpha H'(t,y)\Delta z. \tag{3.11}$$

It follows from the Mean Value Theorem that

$$g_{(t,y)}(\alpha) = \alpha\int_0^1 \left[H'(t+\theta\alpha\Delta t, y+\theta\alpha\Delta y) - H'(t,y)\right]\Delta z\, d\theta.$$

Since $H'(\cdot)$ is uniformly continuous on $\mathcal{N}(\tilde{t},\tilde{y})$ and $\Delta z \to \Delta\tilde{z} := (\Delta\tilde{t}, \Delta\tilde{y})$ as $(t,y)\to(\tilde{t},\tilde{y})$, it follows that for all $(t,y)\in\mathcal{N}(\tilde{t},\tilde{y})$,

$$\lim_{\alpha\downarrow 0}\|g_{(t,y)}(\alpha)\|/\alpha = 0.$$

Then, from (3.10), (3.11) and the facts that $\beta(t,y) \le \gamma\,(\psi(t,y))^{\frac{1}{2}}$ and $\|r\| = \beta(t,y)\bar{t}$, we have that for all $\alpha\in[0,1]$ and all $(t,y)\in\mathcal{N}(\tilde{t},\tilde{y})$,

$$\begin{aligned}
\psi(t+\alpha\Delta t, y+\alpha\Delta y) &= \|H(t+\alpha\Delta t, y+\alpha\Delta y)\|^2\\
&= \|H(t,y) + \alpha H'(t,y)\Delta z + g_{(t,y)}(\alpha)\|^2\\
&= \|(1-\alpha)H(t,y) + \alpha r + g_{(t,y)}(\alpha)\|^2\\
&\le (1-\alpha)^2\psi(t,y) + 2(1-\alpha)\alpha\|H(t,y)\|\|r\| + o(\alpha) + O(\alpha^2)\\
&\le (1-\alpha)^2\psi(t,y) + 4\alpha\|H(t,y)\|\beta(t,y)\bar{t} + o(\alpha) + O(\alpha^2)\\
&\le (1-\alpha)^2\psi(t,y) + 2\alpha\gamma\bar{t}\psi(t,y) + o(\alpha)\\
&\le (1-2\alpha)\psi(t,y) + 2\alpha\gamma\bar{t}\psi(t,y) + o(\alpha)\\
&= [1-2(1-\gamma\bar{t}\ )\alpha]\psi(t,y) + o(\alpha)
\end{aligned}$$

$$\leq [1 - 2\sigma(1 - \gamma\bar{t}\ )\alpha]\psi(t, y) + o(\alpha). \tag{3.12}$$

Thus, by virtue of (3.12), we can find a positive number $\bar{\alpha} \in (0, 1]$ such that (3.8) holds for all $\alpha \in [0, \bar{\alpha}]$ and all $(t, y) \in \mathcal{N}(\tilde{t}, \tilde{y})$. □

Lemma 3.2 *Suppose that $\mathcal{A} \in \mathcal{T}^{m,n}$ is a $\mathcal{Z}$-tensor and $b \in \Re^n_{++}$. Then, Algorithm 1 is well defined at the k-th iteration and $y^k \in \Re^n_{++}$ for any $k \geq 0$. Furthermore,*

$$0 < t^{k+1} \leq t^k \leq \bar{t}, \tag{3.13}$$

and

$$t^k \geq \beta(t^k, y^k)\bar{t}. \tag{3.14}$$

Proof It follows from Proposition 3.1 and Lemma 3.1 that Algorithm 1 is well defined at the k-th iteration and $y^k \in \Re^n_{++}$ for any $k \geq 0$. By the same argument as that given in the proof of [17, Proposition 6], we have that (3.13) and (3.14) hold. □

Lemma 3.3 *If $\mathcal{A} \in \mathcal{T}^{m,n}$ is a $\mathcal{Z}$-tensor and $b \in \Re^n_{++}$, then an infinite sequence $\{(t^k, y^k)\}$ is generated by Algorithm 1. Furthermore, suppose that $(\tilde{t}, \tilde{y})$ is an accumulation point of $\{(t^k, y^k)\}$. Then, $\tilde{y} \in \Re^n_{++}$ and $(\tilde{t}, \tilde{y})$ a solution of $H(t, y) = 0$.*

Proof From Lemma 3.2 and Proposition 3.1, it follows that an infinite sequence $\{(t^k, y^k)\}$ is generated such that $t^k \geq \beta_k\bar{t}$ and $y^k \in \Re^n_{++}$ for all $k \geq 0$. From Algorithm 1, $\psi(t^{k+1}, y^{k+1}) < \psi(t^k, y^k)$ for all $k \geq 0$. In what follows, we denote $\psi_k := \psi(t^k, y^k)$ for notational convenience. Hence, the two sequences $\{\psi_k\}$ and $\{\beta_k\}$ are monotonically decreasing. Since $\psi_k, \beta_k > 0$ $(k \geq 0)$, there exist $\tilde{\psi}, \tilde{\beta} \geq 0$ such that $\psi_k \to \tilde{\psi}$ and $\beta_k \to \tilde{\beta}$ as $k \to \infty$.

Suppose that $(\tilde{t}, \tilde{y})$ is an accumulation point of $\{(t^k, y^k)\}$. Then there exists a subsequence $\{(t^{k_j}, y^{k_j})\}$ of $\{(t^k, y^k)\}$ such that $t^{k_j} \to \tilde{t}$ and $y^{k_j} \to \tilde{y}$ as $k_j \to \infty$. Suppose that $\tilde{y}_i = 0$ for some i. Since $\mathcal{A}$ is a $\mathcal{Z}$-tensor, it follows from (2.4) and (3.1) that $\psi_k \to +\infty$ as $k \to \infty$. This contradiction means that $\tilde{y} > 0$.

Now we show that $(\tilde{t}, \tilde{y})$ is a solution of $H(t, y) = 0$. If the accumulation point $\tilde{\psi}$ of $\{\psi_k\}$ satisfies $\tilde{\psi} = 0$, then from the continuity of $\psi(\cdot)$ and $\beta(\cdot)$, we obtain $\psi(\tilde{t}, \tilde{y}) = 0$ and $\beta(\tilde{t}, \tilde{y}) = 0$. Thus we obtain the desired result. Suppose that $\tilde{\psi} > 0$. By taking a subsequence, if necessary, we may assume that $\{(t^k, y^k)\}$ converges to $(\tilde{t}, \tilde{y})$. It is easy to see from (3.14) that $\tilde{t} \geq \tilde{\beta}\bar{t} > 0$. Then, from Proposition 3.1, $H'(\tilde{t}, \tilde{y})$ exists and is nonsingular. Hence, by Lemma 3.1, there exist a closed neighbourhood $\mathcal{N}(\tilde{t}, \tilde{y})$ of $(\tilde{t}, \tilde{y})$ in $\Re_{++} \times \Re^n_{++}$ and a positive number $\bar{\alpha} \in (0, 1]$ such that for any $(t, y) \in \mathcal{N}(\tilde{t}, \tilde{y})$ and all $\alpha \in [0, \bar{\alpha}]$, we have that $t \in \Re_{++}$, $H'(t, y)$ is invertible and

$$\psi(t + \alpha\Delta t, y + \alpha\Delta y) \leq [1 - 2\sigma(1 - \gamma\bar{t}\ )\alpha]\psi(t, y),$$

where $\Delta t = -t+\beta(t,y)\bar{t}$ and Δy is a solution of the linear system of equations (3.9). Therefore, for a nonnegative integer l such that $\delta^l \in (0,\bar{\alpha}]$, we have $y^k \in \Re^n_{++}$ and

$$\psi(t^k+\delta^l\Delta t^k, y^k+\delta^l\Delta y^k) \le [1-2\sigma(1-\gamma\bar{t}\ )\delta^l]\psi(t^k,y^k)$$

for all sufficiently large k. Then, for every sufficiently large k, $l_k \le l$ and hence $\delta^{l_k} \ge \delta^l$. Thus,

$$\psi(t^{k+1},y^{k+1}) \le [1-2\sigma(1-\gamma\bar{t}\ )\delta^{l_k}]\psi(t^k,y^k) \le [1-2\sigma(1-\gamma\bar{t}\ )\delta^l]\psi(t^k,y^k)$$

for all sufficiently large k. This contradicts the fact that the sequence $\{\psi_k\}$ converges to $\tilde{\psi} > 0$. So, we complete the proof. □

Proposition 3.2 *Let $\mathcal{A} \in \mathcal{T}^{m,n}$ be a $\mathcal{Z}$-tensor and $b \in \Re^n_{++}$. Suppose that $\bar{t}$ and $\tilde{t}$ are two positive numbers such that $\bar{t} \ge \tilde{t} > 0$. Then for any sequence $\{(t^k,y^k)\} \subset \Re_{++}\times\Re^n_{++}$ such that $\tilde{t} \le t^k \le \bar{t}$ and $\|y^k\| \to +\infty$, we have*

$$\lim_{k\to\infty} \psi(t^k,y^k) = +\infty. \tag{3.15}$$

Proof On a contrary, suppose that there exists a sequence $\{(t^k,y^k)\} \subset \Re_{++}\times\Re^n_{++}$ such that $\tilde{t} \le t^k \le \bar{t}$, $\|y^k\| \to +\infty$, and the sequence $\{\psi(t^k,y^k)\}$ is bounded.

We define the index set J by $J := \{i \in [n] : y_i^k \text{ is unbounded}\}$. Then, $J \ne \emptyset$ because $\|y^k\| \to +\infty$. For each k, let

$$\bar{y}_i^k = \begin{cases} y_i^k, & \text{if } i \notin J \\ 1, & \text{if } i \in J \end{cases}, \quad i = 1,2,...,n.$$

Let $\bar{y}^k := (\bar{y}_1^k, \bar{y}_2^k, \dots, \bar{y}_n^k)^\top$. Then, the sequence $\{\bar{y}^k\} \subset \Re^n_{++}$ is bounded. By Lemma 2.2, we can see that $W(y)$ defined in (2.2) is P-function in $\Re^n_{++}$. Hence, for each k, there exists an $i_k \in J$ such that

$$(y_{i_k}^k - 1)\left[W(y^k) - W(\bar{y}^k)\right]_{i_k} > 0.$$

Then, we have

$$\begin{aligned}(y_{i_k}^k - 1)\left[G(t^k,y^k) - G(t^k,\bar{y}^k)\right]_{i_k} &= (y_{i_k}^k - 1)\left[W(y^k) - W(\bar{y}^k) + t^k(y^k-\bar{y}^k)\right]_{i_k} \\ &= (y_{i_k}^k - 1)\left[W(y^k) - W(\bar{y}^k)\right]_{i_k} + t^k(y_{i_k}^k-1)^2 \\ &> t^k(y_{i_k}^k-1)^2.\end{aligned}$$

Since $y_{i_k}^k \to +\infty$ for $i_k \in J$, there exists an integer N such that $y_{i_k}^k > 1$ for all $k \ge N$. So,

$$\left[G(t^k,y^k) - G(t^k,\bar{y}^k)\right]_{i_k} > t^k(y_{i_k}^k - 1).$$

Since $t^k \geq \tilde{t} > 0$ and $y^k_{i_k} \to +\infty$, we have $[G(t^k, y^k) - G(t^k, \bar{y}^k)]_{i_k} \to \infty$ as $k \to \infty$. Note that $\{\|G(t^k, \bar{y}^k)\|\}$ is bounded as $\{\bar{y}^k\}$ is bounded. It then follows that $[G(t^k, y^k)]_{i_k} \to +\infty$. Since J has only a finite number of elements, by taking a subsequence if necessary, we may assume that there exists an $i \in J$ such that

$$[G(t^k, y^k)]_i \to +\infty.$$

Thus, by (3.1), the sequence $\{\psi(t^k, y^k) = \|H(t^k, y^k)\|^2\}$ is unbounded. This is a contradiction which shows that this proposition holds. □

For any given $t \in \Re$, define $\phi_t(y) : \Re^n_{++} \to \Re^+$ by

$$\phi_t(y) = \|G(t, y)\|^2. \tag{3.16}$$

Clearly, $\phi_0(y) = \|W(y)\|^2$, and for any fixed $t \in \Re$, ϕ_t is continuously differentiable at $y \in \Re^n_{++}$ with the gradient given by

$$\nabla\phi_t(y) = 2(G'_y(t, y))^\top G(t, y) \tag{3.17}$$

and $G'_y(t, y)$ is nonsingular at any point $(t, y) \in \Re_+ \times \Re^n_{++}$. Moreover, for any $(t, y) \in \Re \times \Re^n_{++}$, we have

$$\psi(t, y) = t^2 + \phi_t(y). \tag{3.18}$$

Lemma 3.4 *Let $\boldsymbol{C} \subset \Re^n_{++}$ be a compact set. Then for any $\varsigma > 0$, there exists a $\bar{t} \in \Re_{++}$ such that*

$$|\phi_t(y) - \phi_0(y)| \leq \varsigma$$

for all $y \in \boldsymbol{C}$ and all $t \in [0, \bar{t}\,]$.

Proof Since $G(t, y)$ is continuously differentiable at any $(t, y) \in \Re \times \Re^n_{++}$, we can easily show that this lemma holds. □

Theorem 3.1 *Let $f : \Re^n_{++} \to \Re$ be continuously differentiable and coercive, i.e.,*

$$\lim_{\|x\| \to \infty} f(x) = +\infty.$$

Let $\boldsymbol{C} \subset \Re^n_{++}$ be a nonempty and compact set and define ξ to be the least value of f on the (compact) boundary of $\boldsymbol{C}$:

$$\xi := \min_{x \in \partial \boldsymbol{C}} f(x).$$

Assume further that there are two points $\boldsymbol{a} \in \boldsymbol{C}$ and $\boldsymbol{d} \in \Re^n_{++} \backslash \boldsymbol{C}$ such that $f(\boldsymbol{a}) < \xi$ and $f(\boldsymbol{d}) < \xi$. Then there exists a point $\boldsymbol{c} \in \Re^n_{++}$ such that $\nabla f(\boldsymbol{c}) = 0$ and $f(\boldsymbol{c}) \geq \xi$.

This theorem is from [7, Theorem 5.3] by changing the domain of f from $\Re^n$ into $\Re^n_{++}$ and it is easy to prove that this theorem holds. We can use it to prove the following theorem. The proof is similar to that of [14, Theorem 4.6] and [7, Theorem 5.3].

Theorem 3.2 *Suppose that $\mathcal{A} \in \mathcal{T}^{m,n}$ is a nonsingular $\mathcal{M}$-tensor and $b \in \Re^n_{++}$. The sequence $\{(t^k, y^k)\}$ is generated by Algorithm 1. Then,*

(i) $\lim_{k\to+\infty} H(t^k, y^k) = 0$ and $\lim_{k\to+\infty} t^k = 0$.
(ii) *the sequence $\{(t^k, y^k)\}$ is bounded.*
(iii) $\lim_{k\to+\infty} y^k = y^*$*, where y^* is the unique solution of $W(y) = 0$.*

Proof (i). It follows from Lemma 3.2 that an infinite sequence $\{(t^k, y^k)\}$ is generated such that $t^k \geq \beta(t^k, y^k)\bar{t}$ for all $k \geq 0$. From the design of Algorithm 1, $\psi(t^{k+1}, y^{k+1}) < \psi(t^k, y^k)$ for all $k \geq 0$. Hence the sequences $\{t^k\}$, $\{\psi(t^k, y^k)\}$ and $\{\beta(t^k, y^k)\}$ are monotonically decreasing. Since both $\psi(t^k, y^k) > 0$ and $\beta(t^k, y^k) > 0$ for all $k \geq 0$, there exist $\tilde{\psi} \geq 0$ and $\tilde{\beta} \geq 0$ such that $\psi(t^k, y^k) \to \tilde{\psi}$ and $\beta(t^k, y^k) \to \tilde{\beta}$ as $k \to \infty$, respectively. Suppose that $\tilde{\psi} > 0$. Then, from Lemma 3.2,

$$\lim_{k\to+\infty} t^k = \tilde{t} \geq \tilde{\beta}\bar{t} > 0.$$

By Proposition 3.2, it can be easily seen that the sequence $\{(t^k, y^k)\}$ is bounded. From Lemma 3.3, we have $\tilde{\psi} = 0$. This contradiction shows that $\tilde{\psi} = 0$, i.e.,

$$\lim_{k\to+\infty} H(t^k, y^k) = 0 \quad \text{and} \quad \lim_{k\to+\infty} t^k = 0.$$

(ii). Suppose that the infinite sequence $\{(t^k, y^k)\}$ is not bounded. Then the sequence $\{y^k\}$ is not bounded. Let $y^* \in \Re^n_{++}$ be the unique solution of $W(y) = 0$, i.e., the solution of $\phi_0(y) = 0$. Without loss of generality, assume that $\{\|y^k\|\} \to \infty$. Hence there exists a compact set $\boldsymbol{C} \subset \Re^n_{++}$ with $y^* \in \text{int}\boldsymbol{C}$ and

$$y^k \in \Re^n_{++} \backslash \boldsymbol{C} \tag{3.19}$$

for all k sufficiently large. Since

$$\bar{\xi} := \min_{y\in\partial\boldsymbol{C}} \phi_0(y) > 0,$$

we can apply Lemma 3.4 with $\zeta := \bar{\xi}/4$ and conclude that

$$\phi_{t^k}(y^*) \leq \frac{1}{4}\bar{\xi} \tag{3.20}$$

and

$$\xi := \min_{y\in\partial\boldsymbol{C}} \phi_{t^k}(y) \geq \frac{3}{4}\bar{\xi} \tag{3.21}$$

for all k sufficiently large. From Item (i) of this theorem, we have

$$\phi_{t^k}(y^k) \leq \frac{1}{4}\bar{\xi}, \tag{3.22}$$

for all k sufficiently large. Now let us fix an index $\bar{k}$ such that $t^{\bar{k}} \neq 0$ and (3.19)-(3.22) hold. By Proposition 3.2, it is easy to see that for any $\{y^k\}$ with property $\|y^k\| \to +\infty$, we have $\lim_{k\to\infty} \phi_{t^{\bar{k}}}(y^k) = +\infty$. Consequently, by applying Theorem 3.1 with $\boldsymbol{d} := y^k$ and $\boldsymbol{a} := y^*$, we obtain the existence of a vector $\boldsymbol{c} \in \Re^n_{++}$ such that

$$\nabla\phi_{t^{\bar{k}}}(\boldsymbol{c}) = 0 \quad \text{and} \quad \phi_{t^{\bar{k}}}(\boldsymbol{c}) \geq \frac{3}{4}\bar{\xi} > 0.$$

From (3.17) we have $G(t^{\bar{k}}, \boldsymbol{c}) = 0$, i.e., $\phi_{t^{\bar{k}}}(\boldsymbol{c}) = 0$. This contradiction implies that Item (ii) of this theorem holds.

(iii). It follows from Items (i) and (ii) of this theorem that Item (iii) holds immediately. □

Theorem 3.3 *Suppose that $\mathcal{A} \in \mathcal{T}^{m,n}$ is a nonsingular $\mathcal{M}$-tensor and $b \in \Re^n_{++}$. Let $y^* \in \Re^n_{++}$ be the unique solution of (2.3). Then, the sequence $\{(t^k, y^k)\}$ generated by Algorithm 1 converges to $(0, y^*)$ and the convergence rate is Q-quadratic, i.e.,*

$$\|(t^{k+1}, y^{k+1}) - (0, y^*)\| = O\left(\|(t^k, y^k) - (0, y^*)\|^2\right). \tag{3.23}$$

Here, $O(h)$ stands for a function $e : \Re_+ \to \Re_+$, satisfying $e(h) \leq \nu h$ for all $h \in [0, \varrho]$ with some constants $\nu > 0$ and $\varrho > 0$.

Proof First, from Theorem 3.2, the sequence $\{(t^k, y^k)\}$ generated by Algorithm 1 converges to $(0, y^*)$. Now we show that (3.23) holds. In the following, let $z^* := (0, y^*)$, $z^k := (t^k, y^k)$, $\Delta z^k := (\Delta t^k, \Delta y^k)$, and $r_k := (\beta_k \bar{t}, \mathbf{0})$ with $\beta_k := \beta(t^k, y^k)$. From Step 4 of Algorithm 1, we have

$$H(z^k) + H'(z^k)\Delta z^k = r_k \quad \text{and} \quad \|r_k\| = \beta_k \bar{t}. \tag{3.24}$$

Since H is smooth on $\Re \times \Re^n_{++}$ and $H'(z^*)$ is nonsingular, there exist a closed neighbourhood $\mathcal{N}(z^*) \subset \Re \times \Re^n_{++}$ and two scalars L_1 and L_2 such that for all $z := (t, y) \in \mathcal{N}(z^*)$,

$$\|H'(z)^{-1}\| = \|H'(t, y)^{-1}\| \leq L_1$$

and

$$\|H(z) - H(z^*) - H'(z)(z - z^*)\| \leq L_2\|z - z^*\|^2.$$

Then, for z^k sufficiently close to z^*, we have

$$\begin{aligned}\|z^k + \Delta z^k - z^*\| &= \|z^k + H'(z^k)^{-1}[-H(z^k) + r_k] - z^*\| \\ &\leq \|z^k - z^* - H'(z^k)^{-1}H(z^k)\| + L_1\beta_k\bar{t}\end{aligned}$$

$$\begin{aligned}
&\leq L_1\|H(z^k)-H(z^*)-H'(z^k)(z^k-z^*)\|+L_1\beta_k\bar{t}\\
&\leq L_1L_2\|z^k-z^*\|^2+L_1\beta_k\bar{t}\\
&=O(\|z^k-z^*\|^2)+O(\psi_k),
\end{aligned} \tag{3.25}$$

where the last equality follows from the definition of β_k in (3.5), and $\psi_k := \psi(z^k) = \psi(t^k, y^k)$ throughout the proof. Then, because H is smooth at z^*, for all z^k close to z^*,

$$\psi_k = \|H(z^k)\|^2 = O(\|z^k - z^*\|^2). \tag{3.26}$$

Therefore, from (3.25) and (3.26), for all z^k sufficiently close to z^*,

$$\|z^k + \Delta z^k - z^*\| = O(\|z^k - z^*\|^2). \tag{3.27}$$

By (3.27), for any $\epsilon \in (0, \frac{1}{2})$, there is a $k(\epsilon)$ such that for all $k \geq k(\epsilon)$,

$$\|z^k + \Delta z^k - z^*\| \leq \epsilon\|z^k - z^*\|. \tag{3.28}$$

Using (3.24) leads to

$$\begin{aligned}
\|\Delta z^k\| &= \|H'(z^k)^{-1}[-H(z^k)+r_k]\|\\
&\leq L_1\|H(z^k)\|+\bar{t}L_1\psi_k^{\frac{1}{2}}\\
&=(1+\bar{t}\,)L_1\|H(z^k)\|.
\end{aligned} \tag{3.29}$$

It then follows from (3.28) and (3.29) that

$$\begin{aligned}
\|z^k - z^*\| &= \|\Delta z^k\| + \|z^k + \Delta z^k - z^*\|\\
&\leq (1+\bar{t}\,)L_1\|H(z^k)\| + \epsilon\|z^k - z^*\|.
\end{aligned}$$

Consequently, it is clear from $\epsilon \in (0, \frac{1}{2})$ that

$$\|z^k - z^*\| \leq 2(1+\bar{t}\,)L_1\|H(z^k)\|.$$

Since H is smooth at z^*, for all z^k sufficiently close to z^*, we have

$$\begin{aligned}
\psi(z^k+\Delta z^k) &= \|H(z^k+\Delta z^k)\|^2\\
&= O(\|z^k+\Delta z^k - z^*\|^2)\\
&= O(\|z^k - z^*\|^4)\\
&= O(\|H(z^k)-H(z^*)\|^4)\\
&= O(\psi_k^2).
\end{aligned}$$

Therefore, for all z^k sufficiently close to z^* we have

$$z^{k+1} = z^k + \Delta z^k.$$

Hence, by (3.27), we immediately have that (3.23) holds. □

4 Numerical experiments

We have proven that Algorithm 1 (denoted by 'QCA') is globally and quadratically convergent to the solution of problem (1.5). In this section, we further highlight its promising numerical behaviors by solving a synthetic example conducted in the literature.

We wrote the code of QCA in MATLAB 2014a and conducted the experiments on a DELL workstation computer with Intel(R) Xeon(R) CPU E5-2680 v3 @2.5GHz and 128G RAM running on Windows 7 Home Premium operating system. The code of the *homotopy method* proposed by Han [9] was downloaded from Han's homepage[1]. Additionally, we employed the publicly shared tensor toolbox [1] to compute tensor-vector products and semi-symmetrize tensors.

The synthetic example tested in this section comes from [9]. Specifically, we randomly generated a nonnegative tensor $\mathcal{B} := (b_{i_1 i_2 \dots i_m}) \in \mathcal{T}^{m,n}$, whose entries are uniformly distributed in $(0,1)$. To keep $\mathcal{A}$ of (1.5) being a nonsingular $\mathcal{M}$-tensor, we set

$$s = (1+\epsilon) \cdot \max_{1\le i\le n} \left(\sum_{i_2,\cdots,i_m=1}^{n} b_{i i_2 \dots i_m} \right),$$

for some given $\epsilon > 0$. Then, by taking $\mathcal{A} := s\mathcal{I} - \mathcal{B}$, we can always ensure that $\mathcal{A}$ is always a nonsingular $\mathcal{M}$-tensor (see [4,15]) in accordance with $s > \rho(\mathcal{B})$ and

$$\rho(\mathcal{B}) \le \max_{1\le i\le n} \left(\sum_{i_2,\cdots,i_m=1}^{n} b_{i i_2 \dots i_m} \right).$$

Throughout this section, we set $\epsilon = 0.01$ as [6,9]. For the right side vector $b \in \Re^n_{++}$ of (1.5), we also randomly generated it such that all entries are uniformly distributed in $(0,1)$.

Notice that, in [9], the author suggested solving the scaled system of (1.5) instead of the original one, namely, solving

$$\hat{\mathcal{A}} x^{m-1} = \hat{b} \tag{4.1}$$

instead of directly finding solution to (1.5), where $\hat{\mathcal{A}} := \mathcal{A}/\omega$ and $\hat{b} := b/\omega$ with ω being the largest value among the absolute values of components of $\mathcal{A}$ and the entries of b. Moreover, the code of the Homotopy Method in [9] terminates if the residual of the scaled system satisfies

$$\|\hat{\mathcal{A}} x(1)^{m-1} - \hat{b}\|_2 \le 10^{-12}.$$

[1] http://homepages.umflint.edu/~lxhan/software.html

For the proposed QCA method, we have proved that finding a positive solution to (1.5) amounts to solving $H(t, y) = 0$. Hence, as stated in Algorithm 1 (see step 2), we use

$$\mathsf{err} := \|H(t^k, y^k)\| \le \text{Tol} \tag{4.2}$$

to be the stopping criteria.

To illustrate the reliability and efficiency of our proposed QCA, we also implemented it to solve the scaled system (4.1) and compared with the promising Homotopy Method [9]. Note that solving the linear system of equations (3.6) dominates the main computational task of QCA. Hence, how to solve such a subproblem is extremely important to the QCA. Fortunately, the matrix $G'_y(t, y)$ (see Lemma 2.1 and (3.4)) is always nonsingular, which means that (3.6) is a well-defined linear system. In this section, we consider three ways to find solutions to (3.6). Concretely, thanks to the nonsingularity of $G'_y(t, y)$, the first way is solving (3.6) directly by the 'left matrix divide: \', which is roughly the same as the multiplication of the inverse of a matrix and a vector, and another two ways are employments of the MATLAB scripts 'pcg' (which refers to the preconditioned conjugate gradients method) and 'linsolve' (which solves the liner system by using LU factorization with partial pivoting), respectively. In our code for the employment of 'pcg', we set its stopping criteria as '$\max\left\{\min(\mathsf{err}, 10^{-4}), 10^{-6}\right\}$'. Correspondingly, we denote the QCA for scaled system (4.1) with the three ways solving subproblem (3.6) by '**QCA_inv_s**', '**QCA_pcg_s**', and '**QCA_lin_s**', respectively. Throughout, all methods shared the same starting point $\hat{b}^{\frac{1}{m-1}}$. For the parameters of QCA, we took $\delta = 0.5$, $\gamma = 0.8$, $\sigma = 0.2$, and $\bar{t} = 2/(5\gamma)$, which are chosen in accordance with the convergence analysis. Additionally, the stopping tolerance Tol in (4.2) is specified as Tol $= 10^{-10}$. Because of the randomness of the generated $\mathcal{A}$ and b, we conducted 15 groups of (m, n) and randomly generated 10 groups of data sets $(\mathcal{A}, b)$ for each pair of (m, n). As summarized in Table 1, we reported the average performance of the methods. Note that 'itr' represents the average number of iterations; 'ls_itr' denotes the average number of line search steps for finding the smallest nonnegative integers in the QCA; 'pcg_itr' refers to the average number of inner iterations by using 'pcg' to solve subproblem (3.6); 'nt_itr' is the total number of Newton correction steps of Homotopy Method; 'time' denotes the computing time (in seconds) to obtain an approximate solution; 'resi' represents the residual $\|\mathcal{A}x^{m-1} - b\|$ of the original system (1.5) at termination.

It can be easily seen from Table 1 that the proposed QCA is competitive to the promising Homotopy Method in [9], an in particular, for the cases $(m, n) = \{(3, 200),\ (3, 400),\ (3, 500),\ (4, 100),\ (4, 150)\}$, the **QCA_pcg_s** outperforms the other three methods in terms of taking much less computing time to get desired approximate solutions. Indeed, the common feature of these cases is that they

Table 1 Numerical comparison between the QCA and Homotopy Method.

(m, n)	QCA_pcg_s	QCA_inv_s	QCA_lin_s	Homotopy M_s [9]
	itr / ls_itr(pcg_itr) / time / resi[a]	itr / ls_itr / time / resi[a]	itr / ls_itr / time / resi[a]	itr / nt_itr / time / resi[a]
(3, 50)	9.0 / 16.9(108.1) / 0.04 / 8.19	8.8 / 16.4 / 0.04 / 30.8	8.8 / 16.4 / 0.03 / 30.8	5.0 / 10.2 / 0.05 / 21.9
(3, 100)	9.7 / 21.2(115.4) / 1.44 / 4.02	9.6 / 20.9 / 1.17 / 3.36	9.6 / 20.9 / 1.49 / 3.36	5.2 / 9.8 / 1.72 / 0.02
(3, 200)	10.4 / 22.4(116.9) / 0.36 / 0.84	10.4 / 24.1 / 6.95 / 36.5	10.4 / 24.1 / 6.64 / 36.5	5.0 / 8.8 / 6.98 / 3.34
(3, 400)	11.5 / 28.4(125.5) / 2.35 / 0.47	12.3 / 35.9 / 49.99 / 30.8	12.3 / 35.9 / 50.06 / 30.8	5.0 / 7.2 / 39.54 / 335
(3, 500)	11.3 / 26.3(121.0) / 4.39 / 0.65	13.6 / 40.5 / 94.51 / 35.7	13.6 / 40.5 / 96.41 / 35.7	5.0 / 7.5 / 70.00 / 1200
(4, 10)	7.6 / 11.1(64.3) / 0.03 / 4.12	7.6 / 11.1 / 0.03 / 4.57	7.6 / 11.1 / 0.03 / 4.57	5.0 / 10.2 / 0.05 / 5.60
(4, 50)	10.4 / 26.3(112.1) / 0.24 / 13.0	10.4 / 26.3 / 0.24 / 34.9	10.4 / 26.3 / 0.24 / 34.9	5.0 / 8.4 / 0.27 / 34.3
(4, 100)	12.2 / 39.1(142.6) / 3.58 / 45.0	11.8 / 37.2 / 67.73 / 8.14	11.8 / 37.2 / 68.14 / 8.14	5.0 / 8.2 / 64.89 / 528.0
(4, 150)	12.8 / 44.9(153.2) / 18.61 / 1.01	12.7 / 44.6 / 106.50 / 1.27	12.7 / 44.6 / 108.26 / 1.27	5.0 / 8.2 / 108.99 / 1080
(5, 20)	10.6 / 24.2(86.5) / 0.15 / 1.06	10.7 / 24.6 / 0.15 / 1.08	10.7 / 24.6 / 0.14 / 1.08	5.0 / 9.0 / 0.16 / 2.38
(5, 40)	12.5 / 38.6(104.0) / 4.13 / 9.53	11.5 / 35.0 / 3.75 / 9.24	11.5 / 35.0 / 3.75 / 9.24	5.0 / 9.3 / 4.05 / 295.0
(5, 50)	12.9 / 44.0(112.7) / 12.20 / 31.3	13.2 / 44.5 / 12.33 / 0.253	13.2 / 44.5 / 12.40 / 0.253	5.0 / 9.6 / 11.73 / 510.0
(6, 10)	11.1 / 24.8(78.8) / 0.07 / 13.1	10.9 / 23.9 / 0.06 / 39.8	10.9 / 23.9 / 0.07 / 39.8	5.0 / 9.3 / 0.07 / 459.0
(6, 15)	11.7 / 31.4(78.7) / 0.54 / 0.05	11.7 / 31.4 / 0.54 / 13.2	11.7 / 31.4 / 0.54 / 13.2	5.0 / 9.9 / 0.58 / 3510
(6, 20)	13.6 / 42.6(100.1) / 3.22 / 0.69	13.6 / 42.4 / 3.21 / 2.45	13.6 / 42.4 / 3.21 / 2.45	5.0 / 10.0 / 2.99 / 850.0

[a] The resi here denotes the residual in the unit of 10^{-11}, that is, the number 8.19 actually represents 8.19×10^{-11}.

have the relatively higher dimension n. In this situation, the promising efficiency of QCA can be attributed to the efficient solver 'pcg' to the linear system. Additionally, we can also see that both **QCA_inv_s** and **QCA_lin_s** have the almost same performance for each cases. So, we recommend the **QCA_pcg_s** for the first solver to multilinear system (1.5) when it has higher dimensionality.

At the beginning of this section, we mentioned that (4.1) is a scaled system of the original system (1.5) and reported the comparison results by solving such a scaled system. It is noteworthy that the scaled system (4.1) makes both QCA and Homotopy Method have quite different performance, even though both of them share the completely same solution. During the experiments, we observed that the Homotopy method requires much more iterations to get an approximate solution to (1.5) for 3rd-order tensors, and particularly, it is not necessarily convergent for some randomly generated mth-order ($m \geq 4$) cases. So, we also doubt that the proposed QCA would face the same dilemma as the Homotopy Method. To verify the reliability of QCA to the original system (1.5), we also randomly generated 10 groups of $(\mathcal{A}, b)$ and reported the averaged numerical performance in Table 2 with the same settings of parameters used in Table 1, but with the different stopping tolerance $\text{Tol} = 10^{-6}$, where **QCA_inv**, **QCA_pcg**, and **QCA_lin** represent the QCA equipped with different subproblem solvers to the original system (1.5) without scaling technique.

It is clear from the results in Table 2 that the proposed QCA is still powerful to the original system (1.5) without scaling technique. Comparatively, it requires a little more iterations than the case with scaling strategy, thereby taking much more computing time to get an approximate solution. However, it is exciting that applying QCA to the original system directly can get a much more accurate solution. From this perspective, the computational results show that QCA is a quite reliable solver for multilinear systems with $\mathcal{M}$-tensors.

Taking a look back to Algorithm 1, we can see that there are three constants δ, σ, and γ. The reader may be interested in how to choose such parameters in practice. To answer this question, we below investigate the impact (or sensitivity) of the parameters to the proposed QCA equipped with different subproblem solvers. Note that the three parameters only appear in the inequality (3.7). Since $\sigma \in (0, 0.5)$, we can first fix it as $\sigma = 0.2$ empirically. Then, we turn to investigating the impact of δ and γ. More specifically, we tested 8 groups of the parameters $(\delta, \gamma) = \{(0.2, 0.5), (0.2, 0.8), (0.4, 0.5), (0.4, 0.8), (0.5, 0.5), (0.5, 0.8), (0.8, 0.5), (0.8, 0.8)\}$ to investigate the numerical performance. In this part, we only considered the case $(m, n) = (3, 100)$ and generated randomly a pair of $(\mathcal{A}, b)$. The stopping tolerance is specified as $\text{Tol} = 10^{-8}$. We presented six bar graphs in Fig. 1 to illustrated the impact of (δ, γ) to the QCA. All graphs show that the choice of

Table 2 Numerical performance of QCA to the original and scaled systems.

(m,n)	QCA_inv	QCA_pcg	QCA_lin
	itr / ls_itr / time / resi	itr / ls_itr(pcg_itr) / time / resi	itr / ls_itr / time / resi
$(3,200)$	18.9 / 75.4 / 17.42 / 2.19×10^{-13}	17.3 / 61.5(190.0) / 0.60 / 1.95×10^{-13}	18.9 / 75.4 / 12.50 / 2.19×10^{-13}
$(3,400)$	20.9 / 97.4 / 86.84 / 4.01×10^{-13}	19.7 / 77.2(209.2) / 4.17 / 3.88×10^{-13}	20.9 / 97.4 / 87.00 / 4.01×10^{-13}
$(3,500)$	24.6 / 137.9 / 175.76 / 4.36×10^{-13}	19.7 / 81.4(216.1) / 7.76 / 4.79×10^{-13}	24.6 / 137.9 / 175.10 / 4.36×10^{-13}
$(4,50)$	17.7 / 72.9 / 0.36 / 1.17×10^{-13}	18.0 / 75.6(206.7) / 0.39 / 1.22×10^{-13}	17.7 / 72.9 / 0.37 / 1.17×10^{-13}
$(4,100)$	22.7 / 123.9 / 123.01 / 1.55×10^{-13}	20.2 / 102.8(204.5) / 5.17 / 1.51×10^{-13}	22.7 / 123.9 / 123.56 / 1.55×10^{-13}
$(4,150)$	21.9 / 126.8 / 190.31 / 1.96×10^{-13}	22.3 / 127.4(235.2) / 28.58 / 1.91×10^{-13}	21.9 / 126.8 / 192.22 / 1.96×10^{-13}
$(5,40)$	22.5 / 136.5 / 6.52 / 1.12×10^{-13}	22.3 / 131.7(253.5) / 6.56 / 1.04×10^{-13}	22.5 / 136.5 / 6.51 / 1.12×10^{-13}
$(5,50)$	23.8 / 158.4 / 19.91 / 1.08×10^{-13}	24.0 / 157.6(291.8) / 19.99 / 1.16×10^{-13}	23.8 / 158.4 / 19.78 / 1.08×10^{-13}
$(6,15)$	17.9 / 89.7 / 0.75 / 7.17×10^{-14}	18.3 / 91.5(163.7) / 0.78 / 7.09×10^{-14}	17.9 / 89.7 / 0.75 / 7.17×10^{-14}
$(6,20)$	20.1 / 114.4 / 4.20 / 9.58×10^{-14}	20.2 / 117.1(208.2) / 4.26 / 9.44×10^{-14}	20.1 / 114.4 / 4.20 / 9.58×10^{-14}
(m,n)	QCA_inv_s	QCA_pcg_s	QCA_lin_s
	itr / ls_itr / time / resi	itr / ls_itr(pcg_itr) / time / resi	itr / ls_itr / time / resi
$(3,200)$	9.6 / 23.0 / 6.08 / 5.73×10^{-6}	9.8 / 22.2(110.3) / 0.38 / 6.29×10^{-6}	9.6 / 23.0 / 6.11 / 5.73×10^{-6}
$(3,400)$	12.0 / 36.3 / 48.76 / 1.34×10^{-6}	10.0 / 24.9(113.2) / 2.19 / 4.72×10^{-6}	12.0 / 36.3 / 48.29 / 1.34×10^{-6}
$(3,500)$	12.2 / 36.8 / 86.72 / 4.38×10^{-6}	10.6 / 27.4(119.8) / 4.24 / 4.20×10^{-6}	12.2 / 36.8 / 82.99 / 4.38×10^{-6}
$(4,50)$	9.9 / 26.1 / 0.23 / 2.85×10^{-6}	9.9 / 26.6(109.8) / 0.23 / 3.63×10^{-6}	9.9 / 26.1 / 0.23 / 2.85×10^{-6}
$(4,100)$	12.0 / 41.4 / 69.54 / 3.52×10^{-6}	12.1 / 41.0(150.5) / 3.53 / 6.51×10^{-6}	12.0 / 41.4 / 69.22 / 3.52×10^{-6}
$(4,150)$	11.9 / 44.3 / 101.29 / 2.06×10^{-6}	11.9 / 44.3(144.1) / 17.50 / 6.04×10^{-6}	11.9 / 44.3 / 100.40 / 2.06×10^{-6}
$(5,40)$	12.6 / 42.0 / 4.18 / 4.02×10^{-7}	12.6 / 42.0(106.5) / 4.19 / 3.94×10^{-7}	12.6 / 42.0 / 4.17 / 4.02×10^{-7}
$(5,50)$	12.6 / 44.2 / 11.79 / 1.14×10^{-6}	12.7 / 44.4(119.7) / 11.92 / 2.52×10^{-7}	12.6 / 44.2 / 11.78 / 1.14×10^{-6}
$(6,15)$	12.0 / 34.6 / 0.56 / 2.09×10^{-7}	12.0 / 34.6(85.8) / 0.57 / 5.99×10^{-7}	12.0 / 34.6 / 0.56 / 2.09×10^{-7}
$(6,20)$	11.7 / 36.9 / 2.74 / 5.02×10^{-6}	12.3 / 38.9(88.1) / 2.90 / 3.34×10^{-6}	11.7 / 36.9 / 2.75 / 5.02×10^{-6}

(δ, γ) is insensitive to the proposed QCA as long as both of them satisfy $\delta \in (0, 1)$ and $\gamma \in (0, 1)$ such that $\gamma \bar{t} \in (0, 0.5)$.

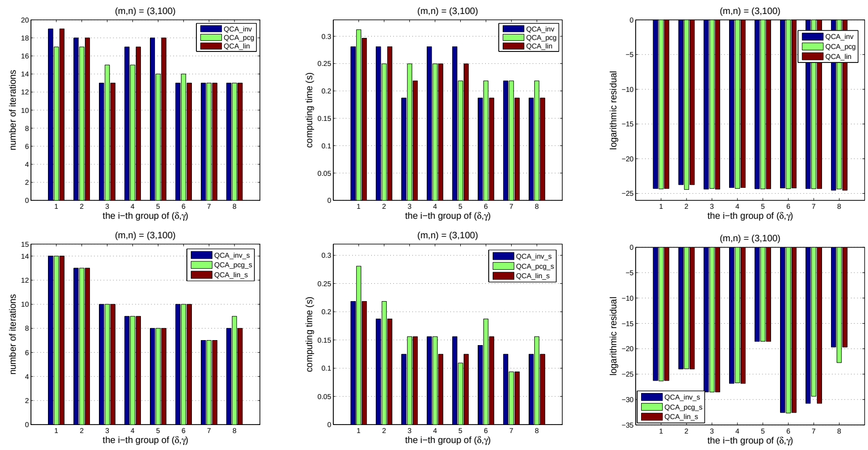

Fig. 1 The impact of parameters (δ, γ) to the numerical performance of QCA with/without the scaling technique. The first row corresponds to the cases of applying QCA to the original system (1.5). The second row corresponds to the case applied to the scaled system (4.1).

Actually, from the above results, we can observe that the proposed QCA attains a high-precise solution by taking about 10 iterations, especially for the scaled system (4.1). The promising performance encourages us to verify the quadratic convergence behaviors intuitively. Thus, we considered two scenarios of (m, n), i.e., $(4, 40)$ and $(5, 30)$, and randomly generated two data sets. In Fig. 2, we plotted the evolution of residual of $\|\mathcal{A}x^{m-1} - b\|$ with respect to the iterations for each scenario, which sufficiently highlights the quadratic convergence as emphasized in the title of this paper.

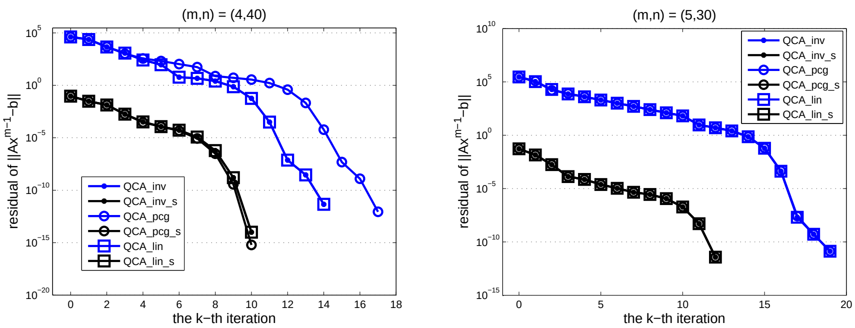

Fig. 2 Evolution of residual of $\|\mathcal{A}x^{m-1} - b\|$ with respect to the iterations. The convergence of QCA is quadratic starting from iterations 8 or 10.

5 Conclusion

In this paper, we have proposed a Newton-type method to solve multilinear systems with $\mathcal{M}$-tensors. In particular, we have shown that the proposed method has a quadratic convergence property. The proposed method can be applied to the following general tensor equation [6]:

$$\mathcal{A}x^{m-1} - \mathcal{B}_{m-1}x^{m-2} - \cdots - \mathcal{B}_2 x = b > 0,$$

where $\mathcal{A} = s\mathcal{I} - \mathcal{B}_m$ is an m-th order nonsingular $\mathcal{M}$-tensor and $\mathcal{B}_p$ is a p-th order nonnegative tensor for $p = 2, 3, ..., (m-1)$. In a similar way to Sections 3 and 4, we can develop Newton-type methods for the above tensor equation and establish their convergence.

Acknowledgements The authors would like to thank Professor Lixing Han for sharing his code on the Homotopy Method. This research of H. He and C. Ling was supported in part by National Natural Science Foundation of China (11571087, 11771113) and the Zhejiang Provincial Natural Science Foundation (LZ14A010003, LY17A010028). L. Qi was supported by the Hong Kong Research Grant Council (Grant Nos. PolyU 502111, 501212, 501913, and 15302114). G. Zhou was supported by the National Natural Science Foundation of China (Grant No. 11601261).

References

1. Bader, B.W., Kolda, T.G., et al.: MATLAB Tensor Toolbox Version 2.6. Available online (2015). URL http://www.sandia.gov/~tgkolda/TensorToolbox/
2. Berman, A., Plemmons, R.: Nonnegative Matrices in the Mathematical Sciences. SIAM, Philadelphia (1994)
3. Brazell, M., Li, N., Navasca, C., Tamon, C.: Solving multilinear systems via tensor inversion. SIAM J. Matrix Anal. Appl. **34**(2), 542–570 (2013)
4. Chang, K., Qi, L., Zhang, T.: A survey on the spectral theory of nonnegative tensors. Numer. Linear Algebra Appl. **20**(6), 891–912 (2013)
5. Ding, W., Qi, L., Wei, Y.: M-tensor and nonsingular m-tensors. Linear Alg. Appl. **439**, 3264–3278 (2013)
6. Ding, W., Wei, Y.: Solving multilinear systems with M-tensors. J. Sci. Comput. **68**, 689–715 (2016)
7. Facchinei, F., Kanzow, C.: Beyond monotonicity in regularization methods for nonlinear complementarity problems. SIAM J. Control Optim. **37**(4), 1150–1161 (1999)
8. Facchinei, F., Pang, J.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer, New York (2003)
9. Han, L.: A homotopy method for solving multilinear systems with M-tensors. Appl. Math. Lett. **69**, 49–54 (2017)
10. Kressner, D., Tobler, C.: Krylov subspace methods for linear systems with tensor product structure. SIAM J. Matrix Anal. Appl. **31**(4), 1688–1714 (2010)
11. Li, D., Xie, S., Xu, H.: Splitting methods for tensor equations. Numer. Linear Algebra Appl. (2017). DOI 10.1002/nla.2102. To appear

12. Li, X., Ng, M.: Solving sparse non-negative tensor equations: algorithms and applications. Front. Math. China **10**, 649–680 (2015)
13. Matsuno, Y.: Exact solutions for the nonlinear Klein–Gordon and Liouville equations in four-dimensional Euclidean space. J. Math. Phys. **28**, 2317–2322 (1987)
14. Qi, H.D.: A regularized smoothing Newton method for box constrained variational inequality problems with p_0-functions. SIAM J. Optim. **10**(2), 315–330 (2000)
15. Qi, L.: Eigenvalues of a real supersymmetric tensor. J. Symbolic Comput. **40**(6), 1302–1324 (2005)
16. Qi, L., Luo, Z.: Tensor Analysis: Spectral Theory and Special Tensors. SIAM, Philadelphia (2017)
17. Qi, L., Sun, D., Zhou, G.: A new look at smoothing newton methods for nonlinear complementarity problems and box constrained variational inequalities. Math. Program. **87**(1), 1–35 (2000)
18. Tobler, C.: Low-rank tensor methods for linear systems and eigenvalue problems. Ph.D. thesis, Eidgenssische Technische Hochschule ETH Zrich, No. 20320 (2012)
19. Xie, Z., Jin, X., Wei, Y.: Tensor methods for solving symmetric M-tensor systems. J. Sci. Comput. (2017). DOI 10.1007/s10915-017-0444-5. To appear
20. Zhang, L., Qi, L., Zhou, G.: m-tensors and some applications. SIAM J. Matrix Anal. Appl. **35**(2), 437–452 (2014)
21. Zwillinger, D.: Handbook of Differential Equations, 3rd edn. Academic Press Inc., Boston (1997)