CrossMark

ORIGINAL PAPER

The sparsest solutions to Z-tensor complementarity problems

Ziyan Luo[1] · Liqun Qi[2] · Naihua Xiu[3]

Abstract Finding the sparsest solutions to a tensor complementarity problem is generally NP-hard due to the nonconvexity and noncontinuity of the involved ℓ_0 norm. In this paper, a special type of tensor complementarity problems with Z-tensors has been considered. Under some mild conditions, we show that to pursuit the sparsest solutions is equivalent to solving polynomial programming with a linear objective function. The involved conditions guarantee the desired exact relaxation and also allow to achieve a global optimal solution to the relaxed nonconvex polynomial programming problem. Particularly, in comparison to existing exact relaxation conditions, such as RIP-type ones, our proposed conditions are easy to verify.

✉ Ziyan Luo
  starkeynature@hotmail.com

  Liqun Qi
  liqun.qi@polyu.edu.hk

  Naihua Xiu
  nhxiu@bjtu.edu.cn

[1] State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, People's Repubic of China

[2] Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, People's Repubic of China

[3] Department of Mathematics, School of Science, Beijing Jiaotong University, Beijing, People's Repubic of China

1 Introduction

The classical compressed sensing theory (see the pioneering work in [4,5,15]) has been gradually generalized and the nonlinear compressed sensing theory has attracted more and more attentions inspired by significant real-life applications such as sparse phase retrieval and sparse pulsation frequency detection in Asteroseimology (see [2] and references therein). Among those nonlinear measurements, the polynomial structure has been employed in many applications cases, such as quadratic measurements in sparse signal recovery [27], and nonlinear compressed sensing with polynomial measurements [26]. Besides the nonlinearity structure carried in many real-life applications, some priori information such as the nonnegativity, can be extensively encountered in communications, DNA microarrays, spectroscopy, tomography, network monitoring, and hidden Markov models [16,20,23,28,33]. In this regard, the optimality condition such as the KKT condition, which has been widely used in the optimization community, might be a good surrogate for the involved feasible set. Under some constraint qualifications, the original problem turns out to finding the sparsest solutions to a special nonlinear complementarity problem with polynomial structures. Mathematically, it can be formulated as

$$(PCP) \quad \begin{aligned} &\min \ \|x\|_0 \\ &\text{s.t.} \ \ F(x) \geq 0, \ x \geq 0, \ \langle x, F(x) \rangle = 0, \end{aligned}$$

where $F = (f_1, \ldots, f_n)$ and all f_i's are polynomial functions.

Apparently, when each f_i in the aforementioned model (PCP) is reduced to be affine (i.e., a linear function added by some constant), (PCP) is exactly to find the sparsest solutions to a linear complementarity problem (LCP for short). The sparsest solution to LCP has been studied in [10,32]. However, to our best knowledge, related work on this topic are very limited, partially but essentially due to the complexity caused by the involved non-convex discontinuous objective function, and partially from the nonlinearity generated by the underlying complementarity constraint. For the former difficulty, many relaxation strategies have been explored by using different surrogates for the ℓ_0 norm, such as the convex ℓ_1 norm [4,5,15], the non-convex ℓ_p norm [9,18,37], the reweighted ℓ_1 norm [6,8], and so on. A natural but essential question arises: is it possible to get an exact solution of the original ℓ_0 norm minimization problem by the relaxation counterpart? If so, what kind of properties should the involved data possess? For the linear measurement case, the well-known restricted isometry property (RIP for short) was introduced to guarantee the desired exactness, which has given a great explanation of the popularity of all sorts of random compressed sensing approaches [3,4]. There are some other exact relaxation conditions on the coefficient matrix for linear constraints, such as the null space property [40], the range space property [41], the s-goodness property [22] and so on. Most of these exact relaxation properties are somehow not easy to verify. In [29], a generalized Z-matrix

was introduced. Together with the nonnegativity of the right-hand side observation vector, it serves as an easy-to-check condition for the desired exact relaxation for the linear sparse optimization. For the nonlinear sparse optimization, such as the sparse LCP, the Z-matrix has been employed to guarantee the exact relaxation [32]. Then how about more general polynomial cases? Can we find some Z-type condition to ensure an exact solution from the relaxation problem for the original sparse polynomial complementarity problem? This is our primary goal in this paper.

Recently, hypermatrices, or tensors,[1] as a higher-order generalization of matrices, have been extensively studied [11,12,30,38], which is closely related to polynomials in terms of those coefficients. This allows us to write a polynomial equation system in a condense form with tensors. In this regard, when the constraint function in the aforementioned problem (PCP) takes the form $F(x) = H(x) - b$ with some homogeneous polynomial function $H(x)$ and some vector $b \in \mathbb{R}^n$, then the feasible set can be reformulated as

$$(TCP) \quad x \geq 0, \ \mathcal{A}x^{m-1} - b \geq 0, \ \langle x, \mathcal{A}x^{m-1} - b \rangle = 0,$$

where $m-1$ is the degree of $H(x)$, x^{m-1} is a rank one tensor of order $m-1$ and dimension n with its $(i_1, \ldots, i_{m-1})$th entry $x_{i_1} \cdots x_{i_{m-1}}$, and $\mathcal{A}$ is an mth order n-dimensional tensor consisting of all the coefficients of $H(x)$ by means of $H(x) = \mathcal{A}x^{m-1}$. Here the tensor product $\mathcal{A}x^{m-1}$ is defined as $(\mathcal{A}x^{m-1})_i = \sum_{i_2,\ldots,i_m=1}^{n} a_{ii_2\ldots i_m} x_{i_2} \cdots x_{i_m}$, for all $i = 1, \ldots, n$. Similarly, we can define $\mathcal{A}x^{m-k}$ as

$$(\mathcal{A}x^{m-k})_{i_1\ldots i_k} = \sum_{i_{k+1},\ldots,i_m=1}^{n} a_{i_1\ldots i_k i_{k+1}\ldots i_m} x_{i_{k+1}} \cdots x_{i_m}, \quad \forall i_1, \ldots, i_k = 1, \ldots, n. \tag{1.1}$$

The above (TCP) is the so-called tensor complementarity problem which has been studied in [7,34,35]. In this paper, we will focus on finding the sparsest solutions to a tensor complementarity problem which can be modeled as

$$(P_0) \quad \begin{aligned} &\min \ \|x\|_0 \\ &\text{s.t.} \ \ \mathcal{A}x^{m-1} - b \geq 0, \ x \geq 0, \ \langle x, \mathcal{A}x^{m-1} - b \rangle = 0. \end{aligned}$$

Mathematically, problem (P_0) is generally NP-hard due to the objective function $\|x\|_0$. Inspired by the scheme of the most popular convex relaxation, we could get the a linear surrogate $e^T x$ resulting from the nonnegativity constraint of x. But the nonlinearity from the tensor complementarity constraints cannot be easily handled and the existing

---

[1] More precisely, here the tensor should be called the hypermatrix. In physics and mechanics, tensors are geometric objects equipped with a transformation law that details how the components of the tensor respond to a change of basis, while hypermatrices are simply multi-dimensional arrays. However, for most papers in tensor decomposition (e.g., the survey paper [25] in SIAM Review, and the numerical optimization methods in [24] in Math. Programming), and in tensor complementarity problems (e.g., [34]), the word "tensors" is used for multi-dimensional arrays. We will then roughly use the word "tensors" throughout this paper.

exact relaxation conditions are not that appropriate since most of them are customized for linear systems.

In this paper, by employing Z-tensors (see e.g., [13]) and the least element theory in nonlinear complementarity problems [36], we present that if the involved b is nonnegative and $\mathcal{A}$ is a Z-tensor, then a sparsest solution of the tensor complementarity problem can be achieved by solving the following polynomial programming problem:

$$(P_1) \quad \begin{aligned} &\min \ e^T x \\ &\text{s.t.} \ \ \mathcal{A}x^{m-1} - b = 0, \quad x \geq 0, \end{aligned}$$

where e is the all one vector. In comparison to those existing exact relaxation conditions for general nonlinear sparse optimization problems [2], our conditions on the coefficients of the polynomial functions are easy to check. This is the main contribution of this paper.

The rest of the paper is organized as follows. The concepts of Z-tensor and Z-function are recalled and some useful properties are presented in Sect. 2. The Z-tensor complementarity problem is introduced and discussed in Sect. 3. The exact relaxation theorem is established in Sect. 4. Concluding remarks are drawn in Sect. 5.

For convenience of presentation, the following notations will be used throughout the paper. We use $\mathbb{R}^n$ and $\mathbb{R}^n_+$ to denote the n-dimensional Euclidean space and its nonnegative orthant respectively. $\mathbb{R}^{n \times n}$ is used to denote the space of all real $n \times n$ matrices. $\mathcal{T}_{m,n}$ is used to stand for the set of all real tensors with order m and dimension n. Vectors are denoted by lowercase letters such as x, matrices are written as capital letters such as A, and tensors are written as calligraphic capital letters such as $\mathcal{A}$.

2 Z-Tensors and Z-functions

As a nonlinear generalization of Z-matrices with non-positive off-diagonal elements, the concept of off-diagonally antitone functions was first introduced by Rheinboldt in [31], which further leads to the definition of Z-functions as stated in [36]. In [21], Isac has redefined the Z-functions equivalently by means of an implication system, which has been also widely used in the community of complementarity problems. In this section, the definition of Z-functions and some useful properties will be recalled. Particularly, to explore the Z-property of homogeneous polynomial functions, the Z-tensor and the partially Z-tensor will be introduced and analyzed.

Definition 1 (*Definition 3.2,* [21]) A mapping $F : \mathbb{R}^n \to \mathbb{R}^n$ is said to be a Z-function if for every $x, y, z \in \mathbb{R}^n_+$ such that $\langle x, y - z \rangle = 0$, we have $\langle x, F(y) - F(z) \rangle \leq 0$.

Proposition 1 (Proposition 3.2, [21]) *A Gâteaux continuous differentiable function* $F : \mathbb{R}^n \to \mathbb{R}^n$ *is a Z-function if and only if for any $x \in \mathbb{R}^n_+$, $\nabla F(x)$ is a Z-matrix.*

Lemma 1 ([21]) *If $F : \mathbb{R}^n \to \mathbb{R}^n$ is a Z-function, then the following implication holds:*

$$x \in \mathbb{R}^n_+, \ y \in \mathbb{R}^n_+, \ \langle x, y \rangle = 0 \Rightarrow \langle x, F(y) - F(0) \rangle \leq 0. \tag{2.1}$$

Moreover, if $F(x) = Ax$ is a linear function, then A is a Z-matrix, which is equivalent to the following implication:

$$x \in \mathbb{R}^n_+, \, y \in \mathbb{R}^n_+, \, \langle x, y \rangle = 0 \Rightarrow \langle x, Ay \rangle \le 0. \tag{2.2}$$

It is known that for any matrix A, Ax is a Z-function if and only if A is a Z matrix, i.e., all of its off-diagonal entries are non-positive. This concept has been extended to the higher order tensors as stated below.

Definition 2 ([39]) Let $\mathcal{A} = (a_{i_1...i_m}) \in \mathcal{T}_{m,n}$. $\mathcal{A}$ is called a Z-tensor if all its off-diagonal entries are nonpositive, i.e., $a_{i_1...i_m} \le 0$ when $\delta_{i_1,...,i_m} = 0$.

Another concept called the partially Z-tensor is introduced here.

Definition 3 Let $\mathcal{A} = (a_{i_1...i_m}) \in \mathcal{T}_{m,n}$. We call $\mathcal{A}$ a partially Z-tensor if for any $i_1 \in [n]$, $a_{i_1 i_2...i_m} \le 0$ for all $i_2, \ldots, i_m$ satisfying $i_1 \notin \{i_2, \ldots, i_m\}$.

Obviously, a Z-tensor is a partially Z-tensor, and both Z- and partially Z-tensors of order $m = 2$ are exactly Z-matrices. Thus, these two concepts both can be regarded as extensions of the Z-matrix. Properties on these two types of tensors are discussed as follows which will play an essential role in the sequel analysis.

Theorem 1 *For any given $\mathcal{A} \in \mathcal{T}_{m,n}$, we have*

(i) *if $\mathcal{A}$ is a partially Z-tensor, then the implication (2.1) holds with $F(x) = \mathcal{A}x^{m-1}$;*
(ii) *if $\mathcal{A}$ is a Z-tensor, then $F(x) = \mathcal{A}x^{m-1}$ is a Z-function.*

Proof (i) Suppose $x, y \in \mathbb{R}^n_+$ with $\langle x, y \rangle = 0$. Easily we can get

$$x_i \ge 0, \, y_i \ge 0, \, x_i y_i = 0, \quad \forall i \in [n]. \tag{2.3}$$

Thus,

$$
\begin{aligned}
&\langle x, F(y) - F(0) \rangle \\
&= \langle x, \mathcal{A}y^{m-1} \rangle = \sum_{i=1}^n x_i (\mathcal{A}y^{m-1})_i \\
&= \sum_{i=1}^n x_i \sum_{i_2,...,i_m=1}^n a_{ii_2...i_m} y_{i_2} \cdots y_{i_m} \\
&= \sum_{i=1}^n \left(\sum_{\substack{i_2,...,i_m=1 \\ i \notin \{i_2,...,i_m\}}}^n a_{ii_2...i_m} x_i y_{i_2} \cdots y_{i_m} + \sum_{\substack{i_2,...,i_m=1 \\ i \in \{i_2,...,i_m\}}}^n a_{ii_2...i_m} x_i y_{i_2} \cdots y_{i_m} \right) \\
&= \sum_{i=1}^n \sum_{\substack{i_2,...,i_m=1 \\ i \notin \{i_2,...,i_m\}}}^n a_{ii_2...i_m} x_i y_{i_2} \cdots y_{i_m} \\
&\le 0,
\end{aligned}
$$

where the last equality follows from (2.3) and the last inequality follows from Definition 3.

(ii) Invoking Proposition 1, it suffices to show that for any $x \in \mathbb{R}_+^n$, $\nabla_x(\mathcal{A}x^{m-1})$ is a Z-matrix. Combining with Lemma 1, we only need to show the implication (2.2) holds with $A := \mathcal{A}x^{m-1}$ for any given $x \in \mathbb{R}_+^n$. Let y, z be any two nonnegative vectors, and $\langle y, z \rangle = 0$. It yields that

$$
\begin{aligned}
& \langle y, \nabla_x(\mathcal{A}x^{m-2})z \rangle \\
&= \sum_{i=1}^{n} y_i \sum_{i_2,\ldots,i_m=1}^{n} \left(a_{ii_2\ldots i_m} + a_{i_2 i\ldots i_m} + \ldots + a_{i_2\ldots i_m i}\right) x_{i_2} \cdots x_{i_{m-1}} z_{i_m} \\
&= \sum_{i=1}^{n} \left(\sum_{\substack{i_2,\ldots,i_m=1 \\ i_m \neq i}}^{n} \left(a_{ii_2\ldots i_m} + a_{i_2 i\ldots i_m} + \ldots + a_{i_2\ldots i_m i}\right) x_{i_2} \cdots x_{i_{m-1}} z_{i_m} y_i \right) \\
& \leq 0,
\end{aligned}
$$

where the second equality is from the complementarity of y and z, and the last inequality is from the fact that $\mathcal{A}$ is a Z-tensor. Thus $\mathcal{F}_\mathcal{A}$ is a Z-function. □

3 Z-Tensor complementarity problems

It is known that a tensor complementarity problem always takes the form

$$
(TCP(\mathcal{A}, b)) \quad x \geq 0, \ \mathcal{A}x^{m-1} - b \geq 0, \ \langle x, \ \mathcal{A}x^{m-1} - b \rangle = 0,
$$

which is actually a special nonlinear complementarity problem. When the involved tensor $\mathcal{A}$ is a Z-tensor, the corresponding $(TCP(\mathcal{A}, b))$ is called a Z-tensor complementarity problem. In this section, we will concentrate on exploiting the properties of such a special class of tensor complementarity problems. We start with recalling a nice property possessed by general nonlinear complementarity problems with Z-functions.

Theorem 2 (Ex. 3.7.21, [17]) *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous Z-function. Suppose that the following nonlinear complementarity problem*

$$
(NCP(F)) \quad x \geq 0, F(x) \geq 0, \langle x, F(x) \rangle = 0
$$

is feasible, i.e., $\mathcal{F} := \{x \in \mathbb{R}^n : x \geq 0, F(x) \geq 0\} \neq \emptyset$. Then $\mathcal{F}$ has a unique least element x^* which is also a solution to $(NCP(F))$.

Inspired by the relation between Z-tensors and Z-functions, we can easily get the following properties for Z-tensor complementarity problems.

Corollary 1 *Let $\mathcal{A}$ be a Z-tensor and $b \in \mathbb{R}^n$. Suppose that the tensor complementarity problem $(TCP(\mathcal{A}, b))$ is feasible, i.e., $\mathcal{F} := \{x \in \mathbb{R}^n : x \geq 0, \mathcal{A}x^{m-1} - b \geq 0\} \neq \emptyset$. Then $\mathcal{F}$ has a unique least element x^* which is also a solution to $(TCP(\mathcal{A}, b))$.*

Proof Theorem 1 tells us that $\mathcal{A}x^{m-1}$ is a Z-function. Utilizing Proposition 1, it is easy to verify that $\mathcal{A}x^{m-1} - b$ is also a Z-function for any $b \in \mathbb{R}^n$. Thus, the desired result follows directly from Theorem 2. □

With a nonnegative b and a partially Z-tensor $\mathcal{A}$ in $(TCP(\mathcal{A}, b))$, the tensor complementarity problem can be equivalent to a multi-linear equation with nonnegative constraints.

Proposition 2 *Let $\mathcal{A}$ be a partially Z-tensor and $b \in \mathbb{R}^n_+$. The following two systems are equivalent:*

(i) $x \in \mathbb{R}^n_+$, $\mathcal{A}x^{m-1} - b \in \mathbb{R}^n_+$, $\langle x, \mathcal{A}x^{m-1} - b \rangle = 0$;
(ii) $x \in \mathbb{R}^n_+$, $\mathcal{A}x^{m-1} - b = 0$.

Proof Trivially, any solution to system (ii) is a solution to system (i). Let y be any solution to system (i). Since $\mathcal{A}$ is a partially Z-tensor and $b \in \mathbb{R}^n_+$, invoking Theorem 2, it yields that

$$
\begin{aligned}
0 &\geq \langle \mathcal{A}y^{m-1} - b, \mathcal{A}y^{m-1} \rangle \\
&= \langle \mathcal{A}y^{m-1} - b, \mathcal{A}y^{m-1} - b \rangle + \langle \mathcal{A}y^{m-1} - b, b \rangle \\
&\geq \|\mathcal{A}y^{m-1} - b\|_2^2.
\end{aligned}
$$

This indicates that $\mathcal{A}y^{m-1} - b = 0$, which implies that y is a solution to (ii). □

Note that Z-tensors are partially Z-tensors. Thus the results in the above proposition hold for Z-tensors.

Corollary 2 *Let $\mathcal{A}$ be a Z-tensor and $b \in \mathbb{R}^n_+$. The following two systems are equivalent:*

(i) $x \in \mathbb{R}^n_+$, $\mathcal{A}x^{m-1} - b \in \mathbb{R}^n_+$, $\langle x, \mathcal{A}x^{m-1} - b \rangle = 0$;
(ii) $x \in \mathbb{R}^n_+$, $\mathcal{A}x^{m-1} - b = 0$.

Utilizing the aforementioned equivalence, we can characterize the feasibility of $TCP(\mathcal{A}, b)$ in terms of the consistency of the corresponding nonnegative constrained multi-linear equation. Before stating the feasibility, we recall the definition of M-tensors, which form an important subclass of Z-tensors.

Definition 4 ([39]) Let $\mathcal{A} \in \mathcal{T}_{m,n}$ be a Z-tensor with $\mathcal{A} = s\mathcal{I} - \mathcal{B}$, where $\mathcal{I}$ is the identity tensor whose diagonal entries are 1 and others 0, $\mathcal{B}$ is a nonnegative tensor and $s \in \mathbb{R}^n_+$. If $s \geq \rho(\mathcal{B})$, then $\mathcal{A}$ is called an M-tensor. If $s > \rho(\mathcal{B})$, then $\mathcal{A}$ is called a strong M-tensor. Here $\rho(\mathcal{B})$ stands for the spectral radius of $\mathcal{B}$.

Proposition 3 *If $\mathcal{A}$ is a strong M-tensor and $b \in \mathbb{R}^n_+$, then $(TCP(\mathcal{A}, b))$ is feasible.*

Proof Let $\mathcal{A} = s\mathcal{I} - \mathcal{B}$ be a strong M-tensor with $\mathcal{B} \geq 0$ and $s > \rho(\mathcal{B})$. Invoking the equivalence as established in Corollary 2, it suffices to show that that there exists some nonnegative x such that $\mathcal{A}x^{m-1} = b$. Let $T_{s,\mathcal{B},b} : \mathbb{R}^n_+ \to \mathbb{R}^n_+$ be the mapping defined as follows:

$$
T_{s,\mathcal{B},b}(x) := (s^{-1}\mathcal{B}x^{m-1} + s^{-1}b)^{\left[\frac{1}{m-1}\right]}, \quad \forall x \in \mathbb{R}^n,
$$

where $x^{\left[\frac{1}{m-1}\right]}$ is the vector with its ith component $x_i^{\left[\frac{1}{m-1}\right]}$, for all $i = 1, \ldots, n$. Easily, we can find that the required nonnegative solution x is exactly a fixed point of this mapping $T_{s,\mathcal{B},b}$. Besides, since $\mathcal{A}$ is a strong M-tensor, applying Theorem 3 in [13], there always exists a positive z such that $\mathcal{A}z^{m-1} > 0$. Denote

$$\alpha := \min_{i \in [n]} \frac{b_i}{(\mathcal{A}z^{m-1})_i}, \quad \text{and} \quad \beta := \max_{i \in [n]} \frac{b_i}{(\mathcal{A}z^{m-1})_i}.$$

Obviously,

$$0 \leq \alpha \mathcal{A}z^{m-1} \leq b \leq \beta \mathcal{A}z^{m-1}.$$

Set $v := \mathcal{A}\left(\alpha^{\frac{1}{m-1}}z\right)^{m-1}$ and $w := \mathcal{A}\left(\beta^{\frac{1}{m-1}}z\right)^{m-1}$. Therefore,

$$\alpha^{\frac{1}{m-1}}z = T_{s,\mathcal{B},v}\left(\alpha^{\frac{1}{m-1}}z\right) \leq T_{s,\mathcal{B},b}\left(\alpha^{\frac{1}{m-1}}z\right),$$

$$\beta^{\frac{1}{m-1}}z = T_{s,\mathcal{B},w}\left(\alpha^{\frac{1}{m-1}}z\right) \geq T_{s,\mathcal{B},b}\left(\beta^{\frac{1}{m-1}}z\right)$$

Note that $T_{s,\mathcal{B},b}$ is an increasing continuous mapping on $\mathbb{R}_+^n$. By employing the fixed point theorem in [1] (also see Theorem 3.1 in [14]), there exists at least one fixed point x of $T_{s,\mathcal{B},b}$ such that $0 \leq \alpha^{\frac{1}{m-1}}z \leq x \leq \beta^{\frac{1}{m-1}}z$. This completes the proof.   □

4 Exact relaxation

Now we are in a position to establish the exact relaxation theorem for the ℓ_0 norm minimization problem (P_0).

Theorem 3 *Let $\mathcal{A}$ be a Z-tensor and $b \in \mathbb{R}_+^n$. If the problem (P_0) is feasible, then (P_0) has a solution x^* which is also the unique solution to the problem (P_1).*

Proof Invoking Theorem 2, we know that $\mathcal{A}x^{m-1}$ is a Z-function. Together with Proposition 1, it is easy to verify that $\mathcal{A}x^{m-1} - b$ is also a Z-function. Thus, Theorem 2 tells us that there exists a unique least element x^* in $\mathcal{F}$ as defined in Corollary 1, which is also a solution of the tensor complementarity problem. The nonnegativity constraint directly yields that x^* is one of the sparsest solutions of (P_0). Utilizing the equivalence as shown in Corollary 2, x^* is definitely the unique solution of (P_1) by the fact that it should be the least element in $\{x \in \mathbb{R}^n : x \geq 0, \mathcal{A}x^{m-1} = b\}$. This completes the proof.   □

Corollary 3 *Let $\mathcal{A}$ be a strong M-tensor and $b \in \mathbb{R}_+^n$. Then problem (P_1) is uniquely solvable and the unique solution is also an optimal solution to problem (P_0).*

Proof This follows directly from Proposition 3 and Theorem 3.   □

Some simple examples are introduced to illustrate the exact relaxation theorem for pursuing the sparsest solution to strong M-tensor complementarity problems.

Example 1 (i) Let $\mathcal{A} = (a_{i_1 i_2 i_3 i_4}) \in \mathcal{T}_{4,2}$ with $a_{1111} = a_{2222} = 1$, $a_{1112} = -2$, and other entries 0, and $b = (0, 1)^\top$. It is known from [19] that $\mathcal{A}$ is a strong M-tensor and the solution set of the corresponding tensor complementarity problem $(TCP(\mathcal{A}, b))$ is $\{(0, 1)^\top, (2, 1)^\top\}$. By direct calculation, $(0, 1)^\top$ is the unique common solution to both problems (P_0) and (P_1).

(ii) Let $\mathcal{A} = (a_{ijk}) \in \mathcal{T}_{3,n}$ be some strictly diagonally dominated Z-tensor (and hence strong M-tensor as shown in [39]) with $a_{ijj} = 0$ for any $i \neq j$, $i, j \in \{1, \ldots, n\}$, and $b \in \mathbb{R}_+^n$ with $\|b\|_0 = 1$. In this case, it is not hard to verify that $x^* = (0, \ldots, 0, \sqrt{\frac{b_i}{a_{iii}}}, 0, \ldots, 0)^\top$ is the unique solution to problem (P_0), where i is the index such that $b_i \neq 0$. This optimal solution is also the unique solution to problem (P_1) by direct verification or by applying the Guass–Seidel iteration method in [14] for the system $\mathcal{A}x^2 = b$ with the zero vector as the initial point.

Remark 1 The above two cases are easy to handle. For general strong M-tensor complementarity problems with $b \in \mathbb{R}_+^n$, the exact relaxation theorem in Corollary 3 always holds, which allows us to get the sparsest solution by solving the corresponding minimization problem with a linear objective function and homogeneous polynomial equality constraints together with a nonnegative constraint. We now apply the aforementioned Guass–Seidel iteration method in [14] for the multi-linear equation $\mathcal{A}x^{m-1} = b$ with the zero vector as the initial point to get the minimal nonnegative solution x^*, which is exactly the solution to problem (P_1) and hence to problem (P_0). Here we will randomly generate 1000 instances of the strong M-tensor $\mathcal{A} \in \mathcal{T}_{3,100}$ by $\mathcal{A} = \eta \mathcal{I} - \mathcal{B}$ with $\mathcal{B} = (b_{ijk}) = rand(100, 100, 100)$, $b_{ijj} = 0$ for any i, $j \in \{1, \ldots, 100\}$, $\eta = 1.1 \times max(sum(sum(\mathcal{B}, 3), 2))$, and $b = 10 \times rand(100, 1)$. Both $\mathcal{B}$ and b will be treated with certain sparsity. The sparsity of the solution is listed in Table 1.

Here $p(\mathcal{A})$ stands for the percentage of nonzero entries of $\mathcal{A}$, which indicates the sparsity of $\mathcal{A}$, $\bar{x}$ is the truncated vector of the optimal solution to problem (P_1) with components larger than $1e$-4, and $s(\bar{x})$ is the average sparsity of $\bar{x}$ for all tested instances. Observe that the sparsity of the solution greatly relies on the sparsity of the input data $\mathcal{A}$ and b. For example, if $b_i \neq 0$, then $x_i \neq 0$ by the structure of $\mathcal{A}$. That is, $\|x^*\|_0 \geq \|b\|_0$. And the sparsity of $\mathcal{A}$ will also greatly effect the sparsity of the solution as Table 1 shows.

Some extended result on exact relaxation theory is discussed as follows. For any matrix $P = (p_{ij}) \in \mathbb{R}^{n \times n}$, we define a linear operator $P\mathcal{A} : \mathcal{T}_{m,n} \to \mathcal{T}_{m,n}$ as follows:

$$(P\mathcal{A})_{i_1 i_2 \ldots i_m} = \sum_{i=1}^{n} p_{i_1 i} a_{i i_2 \ldots i_m}, \quad \forall \mathcal{A} = (a_{i_1 i_2 \ldots i_m}) \in \mathcal{T}_{m,n}.$$

Table 1 The Gauss–Seidel Iteration Method in [14] for Problem (P_1)

$p(\mathcal{A})$	90 %	20 %	90 %	90 %	90 %	20 %	10 %
$\|b\|_0$	100	100	10	2	1	10	10
$s(\bar{x})$	100	100	95.2	95.1	1	45.9	28.1

This is also treated as a special tensor-matrix product in the tensor community (see e.g., [25]). Evidently, if P is an invertible matrix, then $P\mathcal{T}_{m,n} = \mathcal{T}_{m,n}$. However, this operator cannot preserve the Z-property for Z-tensors. Note that the right-hand side of the multi-linear system $\mathcal{A}x^{m-1} = b$ is actually a condense form of n homogeneous polynomials of degree m with any row tensor $\mathcal{A}_i$ formed by the corresponding coefficients of the ith polynomial. If we change the order of these equations in the multi-linear system, the solution set will not be affected. This observation allows us to generalize the result in Theorem 3. For convenience, we use $\mathbb{P}^Z_{m,n}$ to denote the set of all tensors of order m and dimension n which can be transformed to Z-tensors with some permutation matrices, i.e.,

$$\mathbb{P}^Z_{m,n} := \{\mathcal{A} \in \mathcal{T}_{m,n} : P\mathcal{A} \text{ is a } Z\text{-tensor}, P \text{ is a permutation matrix}\}$$

Corollary 4 *Suppose* $\mathcal{A} \in \mathbb{P}^Z_{m,n}$ *and* $b \in \mathbb{R}^n_+$. *If the problem* (P_0) *is feasible, then* (P_0) *has a unique solution* x^* *which is also the unique solution to the problem* (P).

5 Conclusions

To pursuit the sparsest solutions to a tensor complementarity problem can be formulated as an ℓ_0 norm minimization with tensor complementarity constraints, which is always NP-hard. Based on the properties of Z-tensors, we show that one of the sparsest solutions of the Z-tensor complementarity problem can be achieved in polynomial time by solving a polynomial programming problem with a linear objective function, such as the Gauss-Seidel iteration method proposed in [14]. The involved condition on the input of data is indeed an exact relaxation condition for the original ℓ_0 norm minimization. It is worth mentioning that in comparison to other existing exact relaxation conditions in the community of sparse optimization or compressed sensing, our proposed condition is easy to verify. Including the sparse tensor complementarity problem as a special case, the topic of sparse optimization with general nonlinear complementarity constraints deserves further study.

References

1. Amann, H.: Fixed point equations and nonlinear eigenvalue problems in ordered banach spaces. SIAM Rev. **18**(4), 620–709 (1976)
2. Blumensath, T.: Compressed sensing with nonlinear observations and related nonlinear optimization problems. IEEE Trans. Inf. Theory **59**(6), 3466–3474 (2013)
3. Candes, E.J.: The restricted isometry property and its implications for compressed sensing. Comptes Rendus Math. **346**(9), 589–592 (2008)
4. Candes, E.J., Romberg, J.K., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. Commun. Pure Appl. Math. **59**(8), 1207–1223 (2006)
5. Candes, E.J., Tao, T.: Decoding by linear programming. IEEE Trans. Inf. Theory **51**(12), 4203–4215 (2005)

6. Candes, E.J., Wakin, M.B., Boyd, S.P.: Enhancing sparsity by reweighted ℓ_1 minimization. J. Fourier Anal. Appl. **14**(5–6), 877–905 (2008)
7. Che, M., Qi, L., Wei, Y.: Positive definite tensors to nonlinear complementarity problems. J Optim Theory Appl. **168**, 475–487 (2016)
8. Chen, X., Ge, D., Wang, Z., Ye, Y.: Complexity of unconstrained $l_2 - l_p$ minimization. Math. Program. **143**(1–2), 371–383 (2014)
9. Chen, X., Ng, M.K., Zhang, C.: Non-Lipschitz-regularization and box constrained model for image restoration. IEEE Trans. Image Process. **21**(12), 4709–4721 (2012)
10. Chen, X., Xiang, S.: Sparse solutions of linear complementarity problems. Preprint (2014)
11. Comon, P., Golub, G., Lim, L.H., Mourrain, B.: Symmetric tensors and symmetric tensor rank. SIAM J. Matrix Anal. Appl. **30**(3), 1254–1279 (2008)
12. De Lathauwer, L., De Moor, B., Vandewalle, J.: On the best rank-1 and rank-$(R_1, R_2, \cdots, R_N)$ approximation of higher-order tensors. SIAM J. Matrix Anal. Appl. 21(4), 1324–1342 **(electronic)** (2000)
13. Ding, W., Qi, L., Wei, Y.: *M*-tensors and nonsingular *M*-tensors. Linear Algebra Appl. **439**(10), 3264–3278 (2013)
14. Ding, W., Wei, Y.: Solving multi-linear systems with $\mathcal{M}$-tensors. Preprint (2015)
15. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theory **52**(4), 1289–1306 (2006)
16. Donoho, D.L., Tanner, J.: Precise undersampling theorems. Proc. IEEE **98**(6), 913–924 (2010)
17. Facchinei, F., Pang, J.S.: Finite-dimensional Variational Inequalities and Complementarity Problems. Springer Series in Operations Research, vol. I. Springer-Verlag, New York (2003)
18. Foucart, S., Lai, M.J.: Sparsest solutions of underdetermined linear systems via l_q-minimization for $0 < q \leq 1$. Appl. Comput. Harmon. Anal. **26**(3), 395–407 (2009)
19. Gowda, M.S., Luo, Z., Qi, L., Xiu, N.: *Z*-tensors and complementarity problems. arXiv preprint arXiv:1510.07933 (2015)
20. Harmany, Z.T., Marcia, R.F., Willett, R.M.: Spiral out of convexity: sparsity-regularized algorithms for photon-limited imaging. In: IS&T/SPIE Electronic Imaging, pp. 75330R–75330R. International Society for Optics and Photonics (2010)
21. Isac, G.: Complementarity problems. Springer-Verlag, Berlin (1992)
22. Juditsky, A., Nemirovski, A.: On verifiable sufficient conditions for sparse signal recovery via ℓ_1 minimization. Math. Program. **127**(1), 57–88 (2011)
23. Khajehnejad, M.A., Dimakis, A.G., Xu, W., Hassibi, B.: Sparse recovery of nonnegative signals with minimal expansion. IEEE Trans. Signal Process. **59**(1), 196–208 (2011)
24. Kolda, T.G.: Numerical optimization for symmetric tensor decomposition. Math. Program. **151**(1), 225–248 (2015)
25. Kolda, T.G., Bader, B.W.: Tensor decompositions and applications. SIAM Rev. **51**(3), 455–500 (2009)
26. Lauer, F., Ohlsson, H.: Finding sparse solutions of systems of polynomial equations via group-sparsity optimization. J. Glob. Optim. **62**(2), 319–349 (2015)
27. Li, X., Voroninski, V.: Sparse signal recovery from quadratic measurements via convex programming. SIAM J. Math. Anal. **45**(5), 3019–3033 (2013)
28. Liu, Y.F., Dai, Y.H., Luo, Z.Q.: Joint power and admission control via linear programming deflation. IEEE Trans. Signal Process. **61**(6), 1327–1338 (2013)
29. Luo, Z., Qin, L., Kong, L., Xiu, N.: The nonnegative zero-norm minimization under generalized *Z*-matrix measurement. J. Optim. Theory Appl. **160**(3), 854–864 (2014)
30. Qi, L.: Eigenvalues of a real supersymmetric tensor. J. Symb. Comput. **40**(6), 1302–1324 (2005)
31. Rheinboldt, W.C.: On *M*-functions and their application to nonlinear Gauss–Seidel iterations and to network flows. J. Math. Anal. Appl. **32**, 274–307 (1970)
32. Shang, M., Zhang, C., Xiu, N.: Minimal zero norm solutions of linear complementarity problems. J. Optim. Theory Appl. **163**(3), 795–814 (2014)
33. Sheikh, M.A., Sarvotham, S., Milenkovic, O., Baraniuk, R.G.: Dna array decoding from nonlinear measurements by belief propagation. In: IEEE/SP 14th Workshop on statistical signal processing, 2007. SSP'07, pp. 215–219. IEEE (2007)
34. Song, Y., Qi, L.: Properties of some classes of structured tensors. J. Optim. Theory Appl. **165**(3), 854–873 (2014)
35. Song, Y., Qi, L.: Properties of tensor complementarity problem and some classes of structured tensors. arXiv preprint arXiv:1412.0113 (2014)

36. Tamir, A.: Minimality and complementarity properties associated with Z-functions and M-functions. Math. Program. **7**, 17–31 (1974)

37. Xu, Z., Chang, X., Xu, F., Zhang, H.: regularization: a thresholding representation theory and a fast solver. IEEE Trans. Neural Netw. Learn. Syst. **23**(7), 1013–1027 (2012)

38. Yang, Y., Yang, Q.: Further results for Perron–Frobenius theorem for nonnegative tensors. SIAM J. Matrix Anal. Appl. **31**(5), 2517–2530 (2010)

39. Zhang, L., Qi, L., Zhou, G.: M-tensors and some applications. SIAM J. Matrix Anal. Appl. **35**(2), 437–452 (2014)

40. Zhang, Y.: Theory of compressive sensing via ℓ_1-minimization: a non-RIP analysis and extensions. J. Oper. Res. Soc. China **1**(1), 79–105 (2013)

41. Zhao, Y.B.: RSP-based analysis for sparsest and least ℓ_1-norm solutions to underdetermined linear systems. IEEE Trans. Signal Process. **61**(22), 5777–5788 (2013)