On General Conditions for Uniqueness and Stability of Sparse Tensor Signal Reconstruction

Yuan Tian* and Xin Huang
Software School, Dalian University of Technology, Dalian, 116620, China

*Corresponding author email: tianyuan_ca@dlut.edu.cn

Abstract. This paper deals with a fundamental aspect of the inverse problem of robustly reconstructing sparse tensor signals via convex optimization. The traditional vector signal model is extended to tensor model and tensor-space based convex optimization methods are applied to establish the critical results. In particular, by means of some innovative sub-differential analysis for tensor norms and convex geometric analysis in normed tensor space, sufficient conditions to guarantee uniqueness and stability of sparse tensor signal reconstruction are established. In comparison with most current works based on vector signal model (1-order tensor), these conditions are more general and more applicable to tensor signals. Also will these conditions be helpful for establishing practical algorithms for reconstructing high-order sparse tensor signals which are emerging in various data-intensive intelligent applications.

Keywords: Signal processing, Inverse problem, Signal reconstruction, Tensor signal, Uniqueness, Stability.

1. Introduction
Complicated signal reconstruction problems are emerging in more and more scientific and engineering fields, e.g., in modern data-intensive intelligent applications. In traditional paradigm, signals are modeled as sparse vectors and the reconstruction problem is solved (accurately or approximately) by means of optimization analysis with appropriately regularized objective function. Typical analysis, results and algorithms have been developed in compressive sensing theory[1,2]. However, in more and more data-intensive intelligent applications, vector is not the most appropriate model to represent signals. Firstly, lots of complicated signals have more than one attributes. For example, in highly data-intensive radar system[3], measurements are modeled as $y_i = \sum_{ijk} \Psi_{ij,k}X_{ijk} + e_i$ where the echo signal $X_{ijk}$ has distinct attributes of time-delay(range), frequency-shift(radial speed) and direction. Each of which is at best modeled as components in different subspaces instead of in a single vector space[3]. Secondly, sparsity patterns (one of the most important a priori knowledge for signal reconstruction) of complicated signals are richer and diversified than those of vector signals. For example, in imaging applications there are lots of different sparsity patterns of images, which are two-order tensor and cannot be always accurately modeled as a vector, i.e., tensor of order one. In comparison with traditional models and methods mainly dealing with vector signals, the problem of reconstructing sparse tensor signals are more challenging. So far there are very few works on sparse tensor signal reconstruction. Most of such works only deal with matrix signals which are tensors of order two[4,5,6,7]. Some general conditions are established for robustly reconstructing matrix signals in deterministic setting[4,5] and random setting[5,6,7]. Few of them deal with high-order tensors but only in very special forms, e.g., fully-decomposable tensors[5,8].
The methods used in the above-mentioned researches have significant restrictions. Firstly, matrix and tensor signals are modeled as the set of vectors in order to apply the well-developed formulations and conclusions established for vector signals. However, such models cannot well reflect the actual sparsity patterns in high-order tensors. Secondly, these approaches are hard to be extended to dealing with general tensors, which may be not fully-decomposable and with non-$L_1$ norms. Thirdly, some results and conditions developed in these works are not tight\cite{5,8} since the methods are not specific to tensor structures. Some sufficient/necessary conditions and important tight estimates are obtained in \cite{4}, but only for matrix signals.

A complete research on tensor signal reconstruction needs to address the following questions:

- Sufficient and necessary conditions (uniform or non-uniform) to guarantee the reconstruction’s correctness, stability and robustness.
- How to construct the measurement operators which satisfy the above conditions.
- How to design the algorithms to efficiently implement the tensor signals reconstruction.

The above three parts are logically related but each has its own special technical approach. In this paper we focus on aspects in the first part. Our main contributions are some fundamentally sufficient conditions on the measurement operator in signal reconstruction, guaranteeing the reconstructed sparse tensor signal to be unique and stable. We take the convex optimization approach in deterministic setting. In comparison with typical works in current literatures \cite{5,8,10,11}, our conditions are more general and our method is more suitable to dealing with high-order tensor signals. More specifically, the properties guaranteed by these conditions include the reconstruction deviation robustness in terms of any tensor norm, support stability and non-zero components’ sign stability of the reconstructed tensor signal. Further more, the derived reconstruction deviation has linear convergence rate which is asymptotically optimal. These results are presented in the main theorem with a sketched proof.

2. Foundations

Tensors. Intuitively, tensors are vectors with multi-subscripts. Tensors can be also simply regarded as the generalization of matrices. A vector is just a tensor of order one, a matrix is a tensor of order two, and high order tensors can be generated from low order tensors by tensor product operations, e.g., vector of matrices, matrices of matrices, etc.

Formally, tensors are finite linear combination of multi-vectors’ tensor products. Equivalently, tensor spaces are the linear space generated by tensor products of a set of basis.

A tensor space is a linear inner-product vector space on which there are tensor summation and scalar-tensor product operations. In addition, a high order tensor can be reduced to a low order tensor by the reduction operation with other tensors and low order tensors can produce a high order tensor by tensor product operation, denoted by symbol $\otimes$.

Tensors have more and more important applications in data science, engineering and numeric analysis. For space restriction, we suppose the reader have basic knowledge on tensors which can be obtained from the complete reference \cite{9}.

As vectors, various norms can be defined for tensors. In order to reconstruct tensor signals, as in case of vector signals, selecting an appropriate tensor norm as the regularizer is critically important. As the objective function of the convex optimization programming for signal reconstruction, the tensor norm/regularizer should reflect the sparsity pattern in signal. There is no universal tensor norm which can be used as a regularizer to deal with tensor signal reconstruction problem in all cases. Tensor signals with different sparsity patterns require different norms as the regularizer for reconstructing.

Regularizer. We investigate the reconstruction problem with the injective norm\cite{9} as the regularizer.

An injective norm of a $d$-order tensor $T$, denoted as $\|T\|$, is defined as

$$\|T\| = \max \{ (\phi_1 \otimes \ldots \otimes \phi_d)(T) : \phi_i’s \text{ are linear functions of order one with norm } \|\phi_i\|_{\infty} = 1 \}$$

with $\|\phi_i\|_{\infty}$ being the $l_{\infty}$ norm of $\phi_i$, i.e., the maximum of $\phi_i$’s component’s absolute values.

Injective norm is a suitable regularizer to reflect the sparsity patterns of tensor signals appeared in radar systems and data fusion applications in multi-sensor systems. In case of vector signals, this norm
reduces to $l_1$-norm widely used in compressive sensing[1,2]. In case of matrix signals, it reduces to that used in [4]. It must be pointed out that diversified applications need different norms, some of which will be investigated in our future papers.

In addition to injective norm $||\cdot||$, its dual $||\cdot||^*$ and other norms[9] are also needed in our work:

\[
||T||^* = \max \{<S,T>: \text{for all tensor } S \text{ s.t. } ||S||=1 \} \quad (2)
\]

\[
||T||_m = \max \{ \|(\phi_1\otimes\ldots\otimes\phi_d)(T)\): \phi_i's \text{ are standard linear functional } \}
\]

basis of order one \}

\[
||T||^*_m = \max \{<S,T>: \text{for all tensor } S \text{ s.t. } ||S||_m=1 \} \quad (4)
\]

**Convex Optimizer.** In order to reconstruct $d$-order tensor signal $T$, we follow the convex optimization approach, i.e., solve the convex optimizer with the regularizer $||\cdot||$:

Convex Optimizer: \( \min ||S|| \text{ s.t. } y-\Psi(S)||_0 \leq \rho \).

(5)

where $y$ is a measurement vector in space $R^m$, $||\cdot||_0$ is a vector norm to measure the noise magnitude, e.g., $l_2$-norm, $\rho$ is the maximum noise strength, $\Psi$ is a linear mapping. For the real tensor signal $T$, there is a relationship $y=\Psi(X)+z$ where the real noise satisfies $|z|_0 \leq \rho$.

**Basic Fact.** The following fact is basic which proof is in the full version paper.

**Lemma 1** For $d$-order Tensor $T=\sum t_i(i)\otimes\ldots\otimes t_d(i)$, there is the relationship

\[
\tilde{c}||T|| = \{\Sigma_1(\lambda_i)\xi_i(i)\otimes\ldots\otimes\lambda_d(i)\xi_d(i): \xi_i(i) \text{ in } \tilde{c}||t_i(i)||, \lambda_i(i) \geq 0 \}
\]

for all $i$ and $j$, $\lambda_i(i)+\ldots+\lambda_d(i)=1$ and $\lambda_i(i)=0$ for $j: ||t_i(i)||_1 \leq \max||t_i(i)||_1$ \} \quad (6)

3. Auxiliary Results

Now we investigate foundations for the operator $\Psi$ to guarantee the desired properties on reconstructed tensor signal in noisy, in particular, the properties of uniqueness, stability, sign stability and deviation robustness.

The starting point is an elementary theorem that $T^*=\text{argmin}||T||$ s.t. $y-\Psi(T)||_2 \leq \rho$ iff for some real number (multiplier) $\eta^*>0$, $T^*$ is the solution to the unconstrained optimizer $\min||T||+(1/2)\eta^*||y-\Psi(T)||^2$. For the first step towards our goal, we investigate the unconstrained convex optimizer with some given real number $\eta>0$ (its value is set arbitrarily):

\[\text{Problem UCP: } \min||T||+(1/2)\eta \text{ } ||y-\Psi(T)||^2. \quad (7)\]

The following lemma presents the most important characteristics of the solution to (7), which is the foundation for the main results established in next section.

**Convention.** Given an integer $s \geq 1$, a s-sparsity pattern $S = S_1 \cup \ldots \cup S_e$ is a collection of subsets where each $S_i$ is the set of non-zero components (i.e., support) of the tensor signal in its $j$-th component subspace and $|S_i| \leq s$ for each $j$. $-S$ denotes $S$’s complement. Given a linear mapping $\Psi$, symbol $\Psi^s$ denotes its restriction on $S$, $\Psi^T$ denotes its dual mapping and the pseudo-inverse $(\Psi^s)^t \Psi^T$ is denoted as $\Psi^s^{-1}$, if the associated mapping $\Psi^s \Psi^T$ is a bijection.

**Lemma 2** Given measurement data vector $y$, integer $s \geq 1$, s-sparsity pattern $S$ and the linear mapping $\Psi$ with the following properties:

(1) $\Psi^s(U)$ is non-zero when projected onto anything of its component subspace, if $U \neq \emptyset$;

(2) $\Psi^T \Psi^s$ is bijective;

(3) There holds the inequality

\[\eta \text{ max}\{<\Psi^s \Psi^s^{-1}(y) - y, Z>\}, ||Z||=1\}

\[+ \text{ max}\{<\Psi^s \Psi^s^{-1} R, Z>\}, ||Z||=1, ||R|| \leq 1 \} < 1\]

Let $T^*_S=\text{argmin}_{\text{supp}(T) \text{ in } S}||T||+(1/2)\eta ||y-\Psi(T)||^2$ be the solution to (7) with its support in $S$, i.e.,

\[T^*_S= \text{ argmin } ||T||+(1/2)\eta \text{ } ||y-\Psi_S(T)||^2. \quad (9)\]

Then:

(1) $T^*_S$ is the unique solution to the optimizer (9);
(2) $T^*$ is also the unique solution to the optimizer (7), i.e., the global solution to optimizer (7) is unique and is just $T^*$.

(3) Let $W^* = \Psi^*_S(y)$ and $\beta$ be the $d$-tuple of the tensor component’s indices, then $T^*_{S, \beta}$ is unique and is just $T^*$.

Proof Sketch

Step 1. Under the above conditions on $\Psi$, prove the following equation (proved by direct calculation):

$$N(\Psi^*_S: |.|_HS → |.|) < A_{\text{min}}(\Psi^*_S) \left( 1 - N(\Psi^*_S: |.|^* → |.|) \right),$$

where $|.|_HS$ is the Hilbert-Schmidt norm, $A_{\text{min}}(\Psi^*_S) = \min \{ ||\Psi^*_S(U)||: |U|_HS = 1 \}$ and $I_S$ is the identical operator.

Step 2. For arbitrary tensor norm $|.|_a$ on tensor space, the deviation is bounded linearly by $\rho$:

$$|T^* - T|_a \leq 2\rho N(\Psi^*_S: |.|_HS → |.|)$$

(10)

Step 3. For $W^* = \Psi^*_S(y)$ and $\beta$, prove the following equation (proved by direct calculation):

$$N(\Psi^*_S: |.|_HS → |.|) < A_{\text{min}}(\Psi^*_S) \left( 1 - N(\Psi^*_S: |.|^* → |.|) \right)$$

(11)

Proof Sketch The proof is sketched with details presented in the author’s full version paper.

4. Main Results

Now we investigate tensor signal inverse problem by means of the optimizer (5).

Theorem Let integer $s \geq 1$ and the linear mapping $\Psi$ from tensor space to vector space, which satisfies conditions in the following for arbitrary sparsity pattern $S$:

1. $\Psi^*_S(U)$ is non-zero when projected onto any one of its component subspace, if $U \neq 0$;
2. $\Psi^*_S$ is bijective;
3. $N(\Psi^*_S: \Psi^*_S S^{-1} I_S): |.|_HS → |.|^*$

and is just $T^*$.

4. For arbitrary norm $|.|_a$ on tensor space, the deviation is bounded linearly by $\rho$:

$$|T^* - T|_a \leq 2\rho N(\Psi^*_S: |.|_HS → |.|)$$

(10)

5. Sign Robustness: $sgn(T^*_{\beta}) = sgn(T_{\beta})$ for any $\beta$ of the tensor component’s indices in $R$ such that:

$$|T_{\beta}| > \rho \left( N(\Psi^*_R: |.|_HS → |.|) + N(\Psi^*_R: |.|_HS → |.|^*) N(\Psi^*_R: |.|^* → |.|) \right)$$

(11)

Proof Sketch The proof is sketched with details presented in the author’s full version paper.

Step 1. Under the above conditions on $\Psi$, prove the following three intermediary conclusions with the aid of the former lemmas:

a) As the solution to optimizer (5), $T^*$ is unique and $s$-sparse.

b) Let $W^* = \Psi^*_S(y)$, then $T^*_{\beta} \neq 0$ and $sgn(T^*_{\beta}) = sgn(T_{\beta})$ for all $d$-tuple $\beta$ of the tensor component’s indices in $R$ such that:

$$|W^*_{\beta}| > \rho N(\Psi^*_R: |.|_HS → |.|^*) N(\Psi^*_R: |.|^* → |.|)$$

(12)

c) For arbitrary tensor norm $|.|_a$ one has:

$$|T^* - W^*|_a \leq \rho N(\Psi^*_R: |.|_HS → |.|)$$

(12)

Step 2. Because for the tensor $T$ of s-sparsity pattern $R$ and $y = \Psi(T) + z = \Psi_T(T) + z$, $|z|_2 \leq \rho$, one has the following equation (proved by direct calculation):

$$\Psi^*_R \Psi^*_S(y) - y = (\Psi^*_R \Psi^*_S - I_K)z$$

With the aid of this equation and the above intermediary results, a lengthy but direct argument leads to the inequality

$$\max \{ <\Psi^*_R \Psi^*_S(y) - y, V>: |V| = 1 \}$$

$$< \rho A_{\text{min}}(\Psi^*_R) \left( 1 - N(\Psi^*_R: |.|^* → |.|) \right)$$

On basis of this inequality, it’s straightforward to prove $T^*$’s uniqueness.

Step 3. For $W^* = \Psi^*_S(y)$ (which is naturally of s-sparsity pattern $R$), by the inequality (12) one can get $|T^* - W^*|_a \leq \rho N(\Psi^*_R: |.|_HS → |.|)$ for any tensor norm $|.|_a$. Note that $W^* = \Psi^*_S(y)$ derives that $\Psi^*_R(\Psi^*_R(W^*) - y)$ is the zero tensor so on basis of condition (1) there has $\Psi^*_R(W^*) = y$, therefore $\Psi^*_R(W^*) = y = \Psi(T) + z = \Psi_R(T) + z$, or equivalently $\Psi^*_R(W^*) = \Psi^*_R(\Psi^*_R(T) + \Psi^*_R(z))$, hence:
Step 4. Because $|z|_2 \leq \rho$, one has $|W^* - T|_2 \leq \rho N(\Psi R^*-1 \cdot |.||.|) \leq \rho N(\Psi R \cdot |.||.|)$ for any tensor norm $|.|$. By this inequality and $|T^* - W^*|_2 \leq \rho N(\Psi R^*-1 \cdot |.||.|)$ and some general measure inequalities in convex geometry[12] one can derive the recovery deviation’s upper-bound, i.e., $|T^* - T|_2 \leq 2\rho N(\Psi R^*-1 \cdot |.||.|)$

Step 5. By the reconstruction properties on the solution $T^*$ and the fact of $\text{supp}(T) = R$, one can derive by straightforward calculation the relationship

$$T^* - W^* = -\eta^{-1}(\Psi R^T\Psi R)^{-1}(Q^*)$$

where $Q^*$ is a $d$-order tensor in $\mathcal{C}(T^*)$, i.e.,

$$T^* - W^* = -\eta^{-1}(\Psi R^T\Psi R)^{-1}(Q^*)$$

With the aid of (13) and lemma 1, we can derive

$$T^* - T = \Psi R^*-1(z) - \eta^{-1}(\Psi R^T\Psi R)^{-1}(Q^*)$$

Step 6. Since $\text{sgn}(T_\beta) = \text{sgn}(T^*_\beta)$ iff

$$|T_\beta| - |T^*_\beta| = |\Psi R^*-1(z)_\beta - \eta^{-1}(\Psi R^T\Psi R)^{-1}(Q^*)_\beta|$$

particularly, if $X_\beta$ satisfies $|X_\beta| > \max_{\beta}|\Psi R^*-1(z)_\beta| + \eta^{-1}\max_{\beta}|(\Psi R^T\Psi R)^{-1}(Q^*)_\beta|$ then the inequality are true and therefore $\text{sgn}(X_\beta) = \text{sgn}(T^*_\beta)$, which can be proved by inequality (11).

5. Summary

This paper deals with a fundamental aspect of the inverse problem of robustly reconstructing sparse tensor signals via convex optimization. The mostly used vector signal model is extended to tensor model and tensor-space based convex optimization methods are applied to establish the results. The main contributions are some fundamentally sufficient conditions on the measurement operator in sparse tensor signal reconstruction, guaranteeing the reconstructed sparse tensor signal to be unique and stable. We take the convex optimization approach in deterministic setting. More specifically, the properties guaranteed by these conditions include the reconstruction deviation robustness in terms of any tensor norm, support stability and non-zero components’ sign stability of the reconstructed tensor signal. Further more, the established reconstruction deviation’s bound has linear convergence rate which is asymptotically optimal.

In comparison with most other works in current literatures, these conditions are more general and the methods for analysis are more adaptable to dealing with high-order tensor signals. Also will these conditions be helpful for establishing practical algorithms for reconstructing high-order sparse tensor signals which are emerging in various data-intensive applications. Some other work worthwhile in the future is to investigate the sufficient conditions for other kinds of tensor-norms. As a result, Our future works will be on these related topics.

References

[1] Formin, W., Rammer, M.: Inverse problems in signal processing. Prentice-Hall, New York (2018).
[2] Ohm, S.: High dimensional signal processing. Elsevier, Chicago (2017).
[3] Courant, D., Edelman, N.: High resolution radar and related signal analysis techniques, IEEE Signal Proc. Magaz., 45(11), 85-86(2018).
[4] Yuan, T.: On recovery of sparse matrix signals via solving convex optimization, Communications in Computer and Information Science, 16(7), 189-221 (2020).
[5] Wanier, P., Cohen, L.: Hadamard compressive sensing, IEEE Trans. Image Proc., 33(5), 612-635 (2016).
[6] Young, K., Villars, T.: Low-rank matrix reconstruction based-on incomplete measurements, J. Numeric Analysis, 45(1), 89-102 (2017).
[7] Callaway, P.: Greedy algorithms for structured matrix signal approximation, SIAM J. Numeric Computation, 42(6), 997-1015 (2018).
[8] Davenport, N., Lee B.: Efficient compressive approximation for high-dimensional signals. IEEE Trans. Signal Processing, 68(1), 55-72 (2019).
[9] Lawrence, P.: Tensor calculus with applications in numerical analysis, Birkhauser, Archen (2018).

[10] Pohr, B.: Stochastic geometrical approach to reconstructing sparse signals, Discrete. Math. 35(6), 765-784 (2016).

[11] Lax, D.: Recovering sparse signals based on stochastic linear samples, In: Gibson. P(ed), Modern Theoretical and Computational Harmonic Analysis, Springer, Heidelberg (2016).

[12] Rouche, D.: Iterative approximation with sub-linear interpolations for convex domain, Const. Approx., 36(4), 337-386 (2018).