Fundamental Lower Bounds on Number of Random Measurements for 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 problem of robustly reconstructing sparse tensor signals via convex optimization in random setting. The traditional vector signal model is extended to tensor model and tensor-space based probability analysis methods are applied to establish the critical results. In particular, by means of an innovative tensor width estimation and asymptotic convex geometric analysis, fundamental lower bounds on number of random measurements to guarantee high successful reconstruction probability are established. In comparison with most current works based on vector signal model (1-order tensor), these bounds establish foundations to develop effective algorithms for reconstructing high-order sparse tensor signals which are emerging in various data-intensive intelligent signal processing applications.

Keywords: Signal processing, Tensor signal, Convex optimization, Signal reconstruction, Sparsity, Width.

1. Introduction
Complicated signal reconstruction problems are emerging in more and more scientific and engineering fields, particularly in 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 $y_l = \sum_{ljk} \Phi_{ljk} X_{ljk} + e_l$ where the echo signal $X_{ljk}$ 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 tensors 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[5,8] since the methods are not specific to tensor structures. Some sufficient/necessary conditions and important tight estimates are obtained in [4], but only for matrix signals.

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

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. Some basic results have been established by the author for part one in other papers. In this paper we focus on the second part. More specifically, we deal with sparse tensor signal reconstruction problem in random setting. The main contributions are some fundamental lower bounds on number of linear measurements for robust tensor signal reconstruction in noise. Just as the situation in vector signals reconstruction, so far it is unknown how to construct the measurement operators to satisfy the required conditions in deterministic setting, so we deal with the problem in random setting. We will define the width of the set of tensor signals (as a generalization of width of the set of vector signals) and establish the lower bound estimates of the width in terms of specific sparsity pattern parameters. In comparison with other works, e.g.[4,5,6,7], these bounds are more general and asymptotically tight and our methods are more systematic and suitable to dealing with high-order tensor signals.

2. Foundations

Fundamental notions and basic problems to be solved are presented in this section.

**Tensors** (high order signals). 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 [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[9] as the regularizer. An injective norm of a $d$-order tensor $T$, denoted as $\|T\|$, is defined as

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

(1)

where $|\phi|_{\infty}$ is the $L_\infty$ norm of $\phi$, i.e., the maximum of $\phi$’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.

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

**Convex Optimizer I:** \[
\min ||S|| \quad \text{s.t.} \quad y - \Psi(S) \leq \rho. \quad (2)
\]

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

**Convex Optimizer II:** \[
\min ||S|| \quad \text{s.t.} \quad ||Y - \langle A \otimes S, B \rangle||_a \leq \rho. \quad (3)
\]

In this formulation $Y$ is a tensor of lower order (usually much lower than $d$ in practice), $A$ and $B$ are tensors of order $a$ and $b$ and $<.,.>$ is the reduction operator. $|.|_a$ is the tensor norm to measure the noise magnitude, e.g., $l_2$-norm. For the real tensor signal $T$, there is a relationship $Y = \langle A \otimes S, B \rangle + z$ where the real noise satisfies $|z|_a \leq \rho$.

If operator $\Psi$ is appropriately represented, optimizer I can be regarded as a special case of type II. However, the analysis and conclusions on these two formulations are different in some important aspects, so it is advantageous to formulate and deal with them respectively.

**Basic Concepts in Convex Analysis.** A cone $C$ is a domain in linear space such that for all $t > 0$, $tc \in C$. If $R$ is domain in linear space, its polar dual $R^* := \{T: <S, T> \leq 0 \text{ for all } S \in R\}$.

Let $F(T)$ be a convex function defined on tensor space, then define:

\[ D(F; T) := \{U: F(T + tU) \leq F(T) \text{ for some } t \geq 0\}, \]

\[ \partial D(F; T) := \{S: F(R) \geq F(T) + <R - T, S> \text{ for all } R\}. \]

A well known relationship is $U_{\geq 0} \partial F(T) = D(F; T)^\dagger$.

If $C$ is a cone in linear space $J$ on which $\Psi$ is a linear mapping, define its associated singular index

\[ \lambda_{\min, a}(\Psi; C) := \min \{|\Psi(V)|_a: V \in C \text{ and } |V|_b = 1\} \quad (4) \]

where $|.|_a$ and $|.|_b$ are norms on space $J$.

If $C$ is a cone in normed space $(J, |.|_a)$, define its associated width as

\[ w_a(C) := \text{E}_{\text{G}}[\max\{|<G, V>|: V \in C \text{ and } |V|_b = 1\}] \quad (5) \]

where $G$ is the Gaussian random vector on $J$ and $|.|_b$ is a norm on it.

The following fact is basic which proof is in the author’s full version paper.

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

\[ \partial \|T\| = \{\Sigma \lambda_i(i) \xi_i(i) \otimes \ldots \otimes \lambda_d(i) \xi_d(i): \xi_i(i) \text{ in } \partial|t_i(i)|, \lambda_i(i) \geq 0 \text{ for all } i \text{ and } j, \]

\[ \lambda_1(i) + \ldots + \lambda_d(i) = 1 \text{ and } \lambda_k(i) = 0 \text{ for } j: |t_j(i)| \leq \max_i|t_i(i)|\} \quad (6) \]

3. **Bounds on Number of Measurements for Convex Optimizer I**

In this section we prove fundamental lower bounds of measurements counting $m$ for stable recovering the tensor signal $T$ through the convex optimizer I. In component formalism, $y_i = \langle \Psi_i, T \rangle + z_i$ and $\Psi_i$’s are probabilistically independent stochastic tensor of order-$d$.

**Case 1: Gaussian Measurement.** According to the analysis in[9,10], one of crucial ingredients for signal reconstruction in stochastics setting is the estimation on the upper bound of width $\text{w}(D(\|T\|, T))$’s of tensor signal $T$. Lemma 2 presents the main result.

**Lemma 2** Let $T = \Sigma t_i(i) \otimes \ldots \otimes t_d(i)$ be a $d$-order tensor of $s$-sparse $t_i(i)$ as $n$-dimensional vectors, $r$ be the number of members in set $\{(i,j): |t_j(i)| = \max_i|t_i(i)|\}$ where $|.|_1$ is the $l_1$-norm, then
\[
\omega^2(\mathcal{D}(\| \cdot \|, T)) \leq 1 + n^d - r(n^{d-1} - \text{slog}(Cn^{2d}r^2)).
\]  
(7)

As a result, for \( r = n \) there has
\[
\omega^2(\mathcal{D}(\| \cdot \|, T)) \leq 1 + (2d+2)n\text{slog}(Cn).
\]  
(8)

with \( C_1, C_2 \) being numerical constants.

Remark: \( nds \) is tensor signal \( T \)'s sum sparsity. Inequalities (5) and (6) imply that the tensor signal’s width represented by norm \( \| T \|_1 \) is determined by two pattern characteristics, i.e., \( l_1 \)-flatness \( r \) and component-sparsity \( s \). Complicacy can be suppressed by enlarging \( r \) and reducing \( s \).

Proof Sketch For space restriction, only major steps are presented (details in the full version paper).

Step 1. By using the relationship \( \omega^2(\mathcal{D}(\| \cdot \|, T)) \leq E[\min \{ \| G - T \|_{HS^2}^2 : T \in \mathcal{C} \}] \) with \( G \) being a standard Gaussian stochastic tensor and \( \| \cdot \|_{HS} \) is the tensor’s Hilbert-Schmidt norm, the estimation is converted to estimate the upper bound of \( \omega^2(\mathcal{D}(\| \cdot \|, T)) \) for each \( T \in \mathcal{C} \). The relationship is able to be proved by generalizing the arguments in [10,11]).

Step 2. By letting \( G = \Sigma_i g_i(i) \otimes \ldots \otimes g_i(i) \) with \( g_i(i) \sim \text{N}(0,1) \) and using formula (6), \( U = \Sigma_i \lambda_i(i) \xi_i(i) \otimes \ldots \otimes \lambda_i(i) \xi_i(i) \) where \( \xi_i(i) \in \mathcal{C}(t(i), \lambda_i) \geq 0 \) for all \( i \) and \( j \), \( \lambda_i(i) + \ldots + \lambda_j(i) = 1 \) and \( \lambda_j(i) = 0 \) for \( j \neq i \) in \( \{t(i) : t(i) \leq \max \{t(i) \} \} \), denoting the set of the above \( \xi_i(i) \)'s and \( \lambda_j(i)'s \) as \( \Delta \), after some lengthy calculation one has
\[
\omega^2(\mathcal{D}(\| \cdot \|, T)) \leq E[\min \{ \| G - T \|_{HS^2}^2 : T \in \mathcal{C} \}] \leq \min_{\Delta \in \mathcal{C}} \sum_{j=1}^d \| g_j - \tilde{\xi}_j \|_{L^2}^2 + \sum_{j=1}^d \| g_j \|_{L^2}^2
\]

Step 3. Denote the support of \( t \) as \( N(\bar{t}) \) for each \( j \) (note that \( |N\{ j \}| \leq s \) and \( N(\bar{t}) \) as the complimentary set, there has \( \| g_j - \tilde{\xi}_j \|_{L^2}^2 = \| g_{(N_0)} - \tilde{\xi}_j \|_{L^2}^2 + \| g_{(N_0)} - \tilde{\xi}_j \|_{L^2}^2 \). By calculation in combination with estimating some high-dimensional integrals' asymptotic values, one has \( E[\| g_{(N_0)} - \tilde{\xi}_j \|_{L^2}^2] \leq C_0(\rho^2 + \delta^2) \) with \( c \) being some sufficiently small positive number and \( C_0 \) being a numerical constant.

Step 4. By direct calculation one has \( E[\| g_{(N_0)} - \tilde{\xi}_j \|_{L^2}^2] \leq (s + \epsilon)n \).

Step 5. Combine all the above intermediary results, one has \( \omega^2(\mathcal{D}(\| \cdot \|, T)) \leq n^d - r(n^{d-1} - \text{slog}(C_1n^{2d}r^2)) + o(n) \).

On basis of this estimate, one can obtain the important conclusion:

**Theorem 1** Let \( \Psi \)'s components be independently sampled from standard Gaussian distribution \( \mathcal{N}(0,1) \), \( T \) be the real tensor signal with sparsity parameter \( s \) and flatness parameter \( r = n \), measurement vector \( y = \Psi(T) + \varepsilon \) with the inequality \( |\varepsilon| \leq \rho \), \( T \) be the solution to optimizer I with the above data. If number of \( y \)'s components \( m \geq C((2d+2)ns^{d+4}\text{slog})^2 + 2\rho/\delta + dt \), with \( C \) being a numerical constant, there has \( P(\| T - T \|_{HS} \leq \delta) \geq 1 - \exp(-\delta^2/2d) \), i.e., tensor signal \( T \) is able to be recovered robustly in small deviation \( \| T - T \|_{HS} \) with high probability via the convex optimizer I.

**Proof Sketch** This is proved by using lemma 2, Gaussian measure concentration theorems and asymptotic convex geometry techniques [11,12] (details seen in the complete paper).

**Case 2: Sub-Gaussian Measurement** The above result is actually true even for wider stochastic constructions, e.g., when \( \Psi \) is sampled from sub-Gaussian distribution.

**Theorem 2** Suppose \( T \) to be the real tensor signal as that in theorem 1, \( T \) be the solution to optimizer I, \( y = \Psi(T) + \varepsilon \) where \( \Psi \) is a stochastic Sub-Gaussian tensor with \( \Psi \) \( 2 \)-norm \( \omega \). If the number of \( y \)'s components \( m \geq C_2((d^2+2)^{d+4}\text{slog})^2 + 2\rho/\delta + dt \) with \( C_2 \)'s being numerical constants, there has \( P(\| T - T \|_{HS} \leq \delta \geq 1 - \exp(-\delta^2/2d) \), i.e., \( T \) is able to be recovered stably in small deviation \( \| T - T \|_{HS} \) with high probability.

**Proof Sketch** This theorem is proved by sub-Gaussian measure concentration theorems and asymptotic convex geometry techniques (given in the complete paper).

### 4. Bounds on Number of Measurements for Convex Optimizer II

Here we derive the lower bounds on \( m \) for stable recovering tensor signal \( T \) by solving convex optimizer II with operators \( A \) and \( B \) sampled from any sub-Gaussian distributions.

**Theorem 3** Let operators \( A, B \) independently sampled from Sub-Gaussian distributions with \( y_2 \)-norms \( \omega_A \) and \( \omega_B \), \( T \) represent a tensor signal characterized in theorem 1, \( Y = \langle A \otimes S, B \rangle + Z \) with \( |\|Z\|_{HS}^2 \leq \rho \).
$T^*$ be the solution to optimizer II with the above data. If $m \geq dt + C_\delta / \delta + C_2 \delta^{2d} \omega \lambda_{\text{avg}}(ns)^{5d/4} \log n$ with $C_i$'s being numerical constants, there has $P[|T^*-T|_{HS} \leq \delta] \geq 1 - \exp(-r^2/4d^4)$, i.e., $T$ is recovered in small deviation $|T^*-T|_{HS}$ with nearly 100% probability by means of optimizer II.

**Proof Sketch** This theorem is proved (in the complete paper) by a generalized version of lemma 2, tensor’s hierarchical truncation approximation [9] and similar techniques used for the theorem 2.

### 5. Summary and Future Works

Tensor signals are more and more frequently emerging in various data-intensive applications. This paper deals with a fundamental aspect of the problem of robustly reconstructing sparse tensor signals via convex optimization in random setting. The traditional vector signal model is extended to tensor model and innovative tensor-space based analysis methods are used to deal with the problem. In particular, by means of a generalized tensor width estimation and asymptotic convex geometric analysis, fundamental lower bounds on number of random measurements to guarantee highly successful reconstruction probability are established.

In comparison with other works, our bounds are more general and asymptotically tight, and our methods are more systematic and suitable to dealing with high-order tensor signals (tab. 1).

| Signal model | Vector (1-order tensor) | Matrix (2-order tensor) | General tensor (d-order) |
|--------------|-------------------------|------------------------|-------------------------|
| lower bounds on $m$ (Sub-Gaussian measurements) | $O(s \log(n/s))^{[1, 2]}$ | $O((ns)^2 \log n)^{[4, 12]}$ | $O(d^2(ns)^{5d/4} \log n)$ |

These bounds will be helpful to developing effective algorithms for reconstructing high-order sparse tensor signals. The future work will investigate the similar results for other types of tensor regularizers/sparsity-patterns and develop practical algorithms based on these results.

### 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] Paramarchuk, N.: Properties on sub-Gaussian stochastic operators in convex analysis, J. Math. Statistics, 54(3), 375-382 (2019).