Fundamental Lower Bounds on Number of Random Measurements for Structured Matrix Signal Reconstruction

Yuan Tian* and Xin Huang
Software School, Dalian University of Technology, Dalian, P.R. China

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

Abstract. This paper deals with the problem of robustly reconstructing \( n \)-by-\( n \) structured matrix signal via convex optimization in random setting. The traditional vector signal model is extended to matrix model. By means of a generalized matrix width estimation and sub-differential analysis, fundamental lower bounds on number of random measurements to guarantee high successful reconstruction probability are established. In comparison with most current works, these bounds are tighter and the methods are more suitable to be generalized to dealing with high-order tensor signals.

Keywords: Compressive sensing; Matrix signal; Convex optimization; Width.

1. Introduction

Compressive sensing theory develops effective methods to reconstruct signals accurately or approximately from accurate or noisy measurements by exploiting a priori knowledge about the signal\[1,2\]. So far in most works the signals are modeled as vectors of high ambient dimension. However, there are lots of applications in which signals are matrices or even tensors of high orders, particularly in big data processing applications. For example, in highly data-intensive radar system\[3\], measurements are modeled as
\[ y_{kl} = \sum_{ij} \Phi_{kl,ij} X_{ij} + e_{kl} \]
where each \( y_{kl} \) is the echo sampled at specific time \( k \) and specific receiver element \( l \) in a linear or planar array; \( \Phi_{kl,ij} \) is the coefficient of a linear processor; \( e_{kl} \) is the intensity of noise and clutter; signal \( X_{ij} \) is the scattering intensity of a target detected in state cell (\( i,j \)), e.g., a target at specific distance and radial speed, or at specific distance and direction, etc. In applications related to signal sampling/reconstruction, multivariable functions in a linear space spanned by given basis, e.g., \( \{ \psi_\mu(u)\phi_\nu(v) \}_{\mu,\nu} \) are sampled as
\[ s(u,v) = \sum_{\mu,\nu} \psi_\mu(u)\phi_\nu(v) \chi_{\mu,\nu} \]
where \( \chi_{\mu,\nu} \) are the signal’s Fourier coefficients to be reconstructed from the samples. These are typical problems of matrix signals reconstruction and many of them can be naturally extended to even more general tensor signal models.

In comparision with traditional compressive sensing methods which mainly deal with vector signals, the problem of reconstructing matrix, or more generally, tensor signals are more challenging. One difficulty is that such signals have richer and more complicated structures than vector signals. When solving the reconstruction problem via convex programming, it is important to select the appropriate matrix norm (regularizer) for specific signal structure. For example, \( L_1 \)-norm is suitable for general sparsity, nuclear norm is suitable for singular-value-sparsity, and other regularizers are needed for more special or more fine-grained structures, e.g., column-wise sparsity, row-wise sparsity or some hybrid structure. Appropriate regularizer determines the reconstruction’s performance.

So far the works on matrix or tensor signal reconstruction are relatively few, among which typical works include low-rank matrix recovery\[1,4\], matrix completion, Kronecker compressive sensing\[5,6\], etc. Low-rank matrix recovery deals with how to reconstruct the matrix signal with sparse singular values.
from linear measurements using nuclear norm (sum of singular values) as the regularizer, *Kronecker* compressive sensing reconstructs the matrix signal from matrix measurements via matrix *L*1-norm
\[ \sum_{ij} |X_{ij}| \] as the regularizer, dealing with the measurement operator in tensor-product form. In this paper we deal with the challenging problem of reconstructing *n*\texttimes*n* matrix signal \( X = (x_1, \ldots, x_n) \) by convex programming. Signal’s structural features in concern are sparsity and flatness, i.e., each column \( x_i \) is a vector of \( s \)-sparsity and all columns have the same \( l_1 \)-norm. Such signals naturally appear in some important applications, e.g., radar waveform space-time analysis, which will be dealt with as an application in subsequent papers. The regularizer to be used is matrix norm \( ||X||_1 := \max_j |x_j| \) where \( |.| \) is the \( l_1 \)-norm on column vector space.

The main contributions in this paper are some powerful lower bounds on number of linear measurements for robust matrix signal reconstruction in noise. We take the convex geometric approach[7,8,9,10] in random measurement setting and one of the critical ingredients in this approach is to estimate the related widths’ bounds in case of Gaussian and non-Gaussian distributions. As will be seen, these bounds are explicitly controlled by signal’s structural parameters \( r \) and \( s \) which determine matrix signal’s column-wise sparsity and \( l_1 \)-column-flatness respectively. In comparison with most current works, e.g.[3,5], these bounds are tighter and therefore more efficient in practices and our methods are more systematic and suitable to be generalized to dealing with high-order tensor signals.

2. Foundations

2.1. Conventions and Notations

Any vector \( x \) is regarded as column vector, \( x^T \) denotes its transpose (row vector). For vectors \( x \) and \( y \), \( < x, y > \) denotes the scalar product. For matrices \( X \) and \( Y \), \( < X, Y > \) denotes the scalar product \( tr(X^T Y) \). In particular, the *Frobenius* norm \( ||X||_F \) is denoted as \( ||X||_F \). For a positive integer \( s \), \( \Sigma(n,s) \) denotes the set of \( n \times n \) matrices which columns are all of sparsity \( s \), i.e., the number of non-zero components of each column vector is at most \( s \). Let \( S = S_1 \cup \ldots \cup S_n \) be a subset of \( \{ ij \} \) where \( ij = 1, \ldots, n \) and its cardinality \( |S| \leq s \). \( \Sigma(n,s;S) \) denotes the set of \( n \times n \) matrices \{ \( M \): Element \( M_{ij} = 0 \) if \( (i,j) \) not in \( S \) \}. \( S \) is called the *matrix signal’s s-sparsity pattern*. For given \( S \), \( \Sigma(n,s;S) \) is a linear space.

A matrix \( M=(m_1, \ldots, m_n) \) is called \( l_1 \)-*column-flat* if all its columns’ \( l_1 \)-norms \( |m_j|_1 \) have the same value. If \( X_k \) is a group of random variables and \( p(x) \) is some given probability distribution, then \( X_k \sim iid p(x) \) denotes that all these \( X_k \)’s are identically and independently sampled under this distribution.

2.2. Basic Problems

We investigate the problem of reconstructing \( n \times n \) matrix signal \( X = (x_1, \ldots, x_n) \) with \( s \)-sparse and \( l_1 \)-flat column vectors \( x_1, \ldots, x_n \) (i.e., \( ||X||_1 = ||X_j||_1 \) for all \( j \)) by solving the following convex programming problems. The regularizer is matrix norm \( ||X||_1 := \max_j |x_j| \).

\[
\text{Problem MP}(y, \Phi; \eta) : \quad \inf \|Z\|_1 \quad \text{s.t.} \quad |y - \Phi(Z)|_a \leq \eta. \tag{1}
\]

In this setting \( y \) is a measurement vector in \( R^m \) with some vector norm \( |.|_a \) defined on it, e.g., \( |.|_2 \) being the \( l_2 \)-norm. \( \Phi \) is a linear operator and there is a matrix \( X \) (the real signal) satisfying \( y = \Phi(X) + e \) where \( |e|_a \leq \eta \).

\[
\text{Problem MP}(y, A, B; \eta) : \quad \inf \|Z\|_1 \quad \text{s.t.} \quad |Y - AZB^T|_a \leq \eta. \tag{2}
\]

In this setting \( Y \) is a matrix in space \( R^{m \times n} \) with some matrix norm \( |.|_a \) defined on it, e.g., \( |.|_F \) being the *Frobenius*-norm. \( \Phi_{A,B} : Z \rightarrow AZB^T \) is a linear operator and there is a matrix signal \( X \) satisfying \( Y = AXB^T + E \) and \( |E|_a \leq \eta \). \( A \) and \( B \) are full-rank matrices (of rank \( m \)).

2.3. Related Concepts

For brevity the following basic concepts are only presented in form of vectors, however, their generalization to the form of matrices is straightforward.

A cone \( C \) is a subset in \( R^n \) such that \( tC \) is a subset of \( C \) for any \( t > 0 \). For a subset \( K \) in \( R^n \), its polar dual \( K^* := \{ y : <x,y> \leq 0 \text{ for all } x \in K \} \). \( K^* \) is always a convex cone.
For a proper convex function \( F(x) \), there are two important and related sets:

\[
D(F,x):=\{v:F(x+rv) \leq F(x)\} \text{ for some } t>0, \quad \partial F(x) = \{u:F(y) \geq F(x) + \langle y-x, u \rangle \text{ for all } y\}.
\]

and an important relation is \( D(F,x)^c \) = the union of \( t\partial F(x) \) with all \( t>0 \).

Let \( |\cdot| \) be some vector norm and \( |\cdot|^* \) be its conjugate norm, i.e., \( |u|^*:=\max\{\langle x,u\rangle: |x| \leq 1\} \) (e.g., \( \|X\|_1^*=\sum_j |x_j|_\infty \)) then \( \partial|x| = \{u: |x| = \langle x,u \rangle \text{ and } |u|^* \leq 1\} \).

Let \( K \) be a cone in a vector space \( L \) on which \( \Phi \) is a linear operator, the minimum singular value of \( \Phi \) with respect to \( K \), norm \( \|\cdot\|_\beta \) on \( L \) and norm \( \|\cdot\|_\alpha \) on \( \Phi \)'s image space is defined as

\[
\lambda_{\min,\alpha,\beta}(\Phi;K):=\inf\{\|\Phi u\|_\alpha: u \text{ in } K \text{ and } |u|=1\} \tag{3}
\]

When both \( \|\cdot\|_\beta \) and \( \|\cdot\|_\alpha \) are \( l_2 \) (or Frobenius) norms, \( \lambda_{\min,\alpha,\beta}(\Phi;K) \) is simply denoted as \( \lambda_{\min}(\Phi;K) \).

Let \( K \) be a cone in normed space \((L,\|\cdot\|_\beta)\), its conic Gaussian width is defined as

\[
w_{\beta}(K):=E_{\xi}[\sup\{\langle \xi,u\rangle: u \in K \text{ and } |u|=1\}] \tag{4}
\]

where \( \xi \) is the random vector on \( L \) sampled under standard Gaussian distribution. When \( \|\cdot\|_\beta \) is \( l_2 \) or Frobenius norm on \( L \), \( w_{\beta}(K) \) is simply denoted as \( w(K) \).

Now present an important fact which proof is in the author’s full version preprint [11].

**Lemma 1** For \( n\)-by-\( n \) matrix \( X=(x_1,\ldots,x_n) \), the subdifferential of matrix norm \( \|X\|_1:=\max_j |x_j|_1 \) is \( \partial\|X\|_1=\{(\lambda_1,\xi_1),\ldots,\lambda_n,\xi_n)\}: \xi_j \in \partial|x_j|_1 \text{ and } \lambda_j \geq 0 \text{ for all } j, \lambda_1+\ldots+\lambda_n=1 \text{ and } \lambda_j=0 \text{ for } j: |x_j|_1<\max_j |x_j|_1 \}

3. Bounds on Number of Measurements for Robust Reconstruction via Solving \( MP^{(2)}_{y,\Phi,\eta} \)

Now establish some lower bounds on number of measurements \( m \) for robustly reconstructing the matrix signal \( X \) by solving the convex programming problem \( MP^{(2)}_{y,\Phi,\eta} \), namely \( \inf\|Z\|_1 \text{ s.t } |y-\Phi(Z)|_2 \leq \eta \) where \( y=\Phi(X)+\epsilon, |\epsilon|_2 \leq \eta \), \( \Phi \) is a linear operator. In the equivalent component-wise formulation, \( y_j=\Phi(x_j)+\epsilon_j \) where for each \( j=1,\ldots,m \), \( \Phi \) is a random matrix independent each other and \( \Phi_i \)'s entries are independently sampled under standard Gaussian \( N(0,1) \) or sub-Gaussian distribution.

3.1. Case 1: Gaussian Measurement Operator \( \Phi \)

Based upon the general theory in [7,8], one of the critical steps in random approach is to estimate the width \( w(D(\|\cdot\|_1, X)) \)’s upper bound for matrix signal \( X=(x_1,\ldots,x_n) \) with \( s \)-sparse column vectors \( x_1,\ldots,x_n \). This is done in Lemma 2.

**Lemma 2** Given \( n\)-by-\( n \) matrix \( X=(x_1,\ldots,x_n) \) with \( s \)-sparse column vectors \( x_1,\ldots,x_n \). Let \( r \) (called \( l_1 \)-column-flatness parameter hereafter) be cardinality of the set \( \{j: |x_j|_1=\max_i |x_i|_1\} \), i.e., the number of column vectors which have the maximum \( l_1 \)-norm, then

\[
w^2(D(\|\cdot\|_1, X)) \leq 1+n^2-r(n-s\log(Cn^4r^2)). \tag{5}
\]

In particular, when \( r=n \) then

\[
w^2(D(\|\cdot\|_1, X)) \leq 1+ns\log(Cn^8). \tag{6}
\]

where \( C \) is an absolute constant.

**Remark:** \( ns \) is the total sparsity of the matrix signal \( X \). This lemma indicates that the matrix signal complexity (width) encoded by regularizer \( \|X\|_1 \) is controlled by two structural parameters, the column-sparsity \( s \) and \( l_1 \)-column-flatness \( r \). Complexity gets lower with smaller \( s \) and larger \( r \).

**Proof** We start with the basic inequality \( w^2(D(\|\cdot\|_1, X)) \leq E_G[\inf\|G^{-1/2}v\|_2^2]: t>0, V \in \partial\|X\|_1 \} \) where \( G \) is a random matrix with entries \( G_{ij} \sim \text{ iid } N(0,1) \) (for its proof and generalization, see [7,8]).

Set \( G=(g_1,\ldots,g_m) \) where \( g_i \sim \text{ iid } N(0,1) \). By lemma 1, \( V=(\lambda_1,\xi_1,\ldots,\lambda_n,\xi_n) \) where w.l.o.g. \( \lambda_i \geq 0 \text{ for } j=1,\ldots,r, \lambda_1+\ldots+\lambda_r=1, \lambda_0=0 \text{ for } j>2r+1; |x_j|_1=\max_i |x_i|_1 \text{ for } j=1,\ldots,r \text{ and } |x_j|_1<\max_i |x_i|_1 \text{ for } j>1+r; \xi_i=sgn(X_{ij}) \text{ for } X_{ij} \neq 0 \text{ and } |\xi_i|_1 \leq 1 \text{ for all } i \text{ and } j. \)

Then

\[
w^2(D(\|\cdot\|_1, X)) \leq E_G[\inf_{r=0,\lambda_j,\xi_j \text{ specified as the above}} {\sum_{1 \leq j \leq r} g_j - \lambda_j \xi_j \xi_j^2 + \sum_{1 \leq r+1} g_j^2}]\]
For each \( j=1, \ldots, r \) let \( S(j) \) be the support of \( x_j \) (so \( |S(j)| \leq s \) and \( -S(j) \) be its complimentary set, then \( |g_j - \tilde{\partial}_j \xi_j| = |g_{j-S(j)} - \tilde{\partial}_j \xi_{j-S(j)}| + |g_{j-S(j)} - \tilde{\partial}_j \xi_{j-S(j)}| \). Notice that all components of \( \xi_{j-S(j)} \) are \( \pm 1 \) and all components of \( \xi_{j-S(j)} \) can be any value in the interval \([-1,+1] \). Select \( \lambda_1 = \cdots = \lambda_r = 1/r \), let \( \varepsilon > 0 \) be arbitrarily small positive number and select \( t(\varepsilon) \) such that \( P[|g| > t(\varepsilon)/r] \leq \varepsilon \) where \( g \) is a standard scalar Gaussian random variable (i.e., \( N(0,1) \)) and \( \varepsilon \) can be \( \exp(-t(\varepsilon)^2/2r^2) \). For each \( j \) and each \( i \) outside \( S(j) \), set \( \xi_j(i) = r g_j(i)/t(\varepsilon) \) if \( |g_j(i)| \leq t(\varepsilon)/r \) (in this case \( g_j(i) - \tilde{\partial}_j \xi_j(i) = 0 \) and otherwise \( \xi_j(i) = \text{sgn}(g_j(i)) \) (in this case \( g_j(i) - \tilde{\partial}_j \xi_j(i) = |g_j(i)| - t(\varepsilon)/r \)), then \( g_{j-S(j)} - \tilde{\partial}_j \xi_{j-S(j)} = 0 \) when \( g_{j-S(j)} \). \( \varepsilon < t(\varepsilon)/r \), hence:

\[
E[|g_{j-S(j)} - \tilde{\partial}_j \xi_{j-S(j)}|^2] = \int_{\omega=0} \int_{\omega=0} du du | |g_{j-S(j)} - \tilde{\partial}_j \xi_{j-S(j)}|^2| = 2 \int_{\omega=0} \int_{\omega=0} du du | |g_{j-S(j)} - \tilde{\partial}_j \xi_{j-S(j)}| > u| \leq 2 \int_{\omega=0} \int_{\omega=0} du du \text{w}^2(\|g\|/r > (n-s)^{1/2} u) \leq 2(2-s) \int_{\omega=0} \int_{\omega=0} du \text{w}^2(\|g\| > (n-s)^{1/2} u)) \leq 2(2-s) \int_{\omega=0} \int_{\omega=0} du \text{w}^2(\|g\| > (n-s)^{1/2} u)^2/2 \leq C_0(2-s) \text{w}^2(\|g\| > (n-s)^{1/2} u)) \leq C_0(n-s)\text{w}^2(\|g\| > (n-s)^{1/2} u))\text{.}
\]

where \( C_0 \) is an absolute constant. On the other hand:

\[
E_\omega[|g_{j-S(j)} - \tilde{\partial}_j \xi_{j-S(j)}|^2] = E_\omega[|g_{j-S(j)}|^2] + (t(\varepsilon)^2/2r^2) = 1/2 + (t(\varepsilon)^2/2r^2) = 1/2 + 2\log(1/c)\text{.}
\]

Hence \( \text{w}^2(\|g\| > (n-s)^{1/2}) \)

\[
\leq (1+2\log(1/c))r + (n-s)^{1/2} \leq n^2 - (n-s)\text{w}^2(\|g\| > (n-s)^{1/2} u)^2/2 \leq C_0(n-s)\text{w}^2(\|g\| > (n-s)^{1/2} u))\text{.}
\]

Based on this lemma, we get the following results [details in the full version preprint [11]]:

**Theorem 1** Suppose \( \Phi_{\omega,j} \sim_{\text{ind}} N(0,1) \), let \( X \) be a columnwise \( s \)-sparse and \( l_1 \)-column-flat matrix signal, \( y = \Phi(X) + \epsilon \) where \( \|\epsilon\| \leq \eta \), \( X \) be the minimizer of problem \( MP^{(2)}_{y, \Phi, \eta} \). If the measurement vector \( y \)'s dimension \( m \geq (t + 2\eta/\delta + (n-s)\log(Cn^{\delta}))^{1/2} \), where \( C \) is an absolute constant, then \( P[|X^* - X| < \delta] \geq 1 - \exp(-t^2/2) \), i.e., matrix signal \( X \) can be reconstructed robustly with respect to the error norm \( |X^* - X| \) with high probability by solving convex programming \( MP^{(2)}_{y, \Phi, \eta} \).

**3.2. Case 2: Sub-Gaussian Measurement Operator \( \Phi \)**

The following result can be obtained in a similar way as the above:

**Theorem 2** Let \( X \) and \( X^* \) be respectively the matrix signal and the minimizer of \( MP^{(2)}_{y, \Phi, \eta} \) where the signal \( X \) is column-wise \( s \)-sparse and \( l_1 \)-column-flat, \( y = \Phi(X) + \epsilon \), each \( \Phi_{k} \sim_{\text{ind}} \Phi \) where \( \Phi \) is a random matrix satisfying the canonical conditions with parameters \( a, \rho \) and \( \sigma \). If the measurement vector \( y \)'s dimension

\[
m \geq (C_1 \rho^4/\alpha)(at + 2\eta/\delta + \sigma C_2(\rho^4 ns\log(Cn^{\delta})))^{1/2} \text{.}
\]

where \( C_1 \)'s are absolute constants, then \( P[|X^* - X| < \delta] \geq 1 - \exp(-C_4 t^2) \), i.e., matrix signal \( X \) can be reconstructed robustly with respect to the error norm \( |X^* - X| \) with high probability.

4. **Bounds on Number of Measurements for Robust Reconstruction via Solving \( MP^{(F)}_{y, A, B, \eta} \)**

Now establish useful lower bounds on number of measurements \( m \) for robustly reconstructing the matrix signal \( X \) by convex programming \( MP^{(F)}_{y, A, B, \eta} \), namely \( \inf \|Z\|_1 \text{ s.t. } |Y - AZB|^2 \leq \eta \) where \( Y = AXB + E, \|E\| \leq \eta \). In the canonical-case formulation, \( y = \Phi(X) = \sum_j A_{kj}X_jB_{kj} \) for each \( 1 \leq k, j \leq m, A_{kj} \sim_{\text{ind}} B_{kj} \sim_{\text{ind}} N(0,1) \) or sub-Gaussian distribution, matrix \( A, B \) are independent each other.
4.1. Case 1: And B are both Gaussian
The related width is estimated in Lemma 3 based on which the main result is established in this case. The proof is more involved than in last cases because now the measurement operator is not even sub-Gaussian. Details are presented in preprint[11].

Lemma 3 Given n-by-n matrix $X=(x_1,\ldots,x_n)$ with $s$-sparse column vectors $x_1,\ldots,x_n$, $r$ is cardinality of $\{j: |x_j|_1=\max_i|x_i|_1\}$, $\Phi_{A,B}, \Gamma_X, W(\Gamma_X; \Phi_{A,B})$ are specified as the above and $A_{ki}\sim iid B_{ij}\sim iid N(0,1)$, then

$$W^2(\Gamma_X; \Phi_{A,B}) \leq 1 + n^2 - r(n - \log^2(cn^r)).$$

where $c$ is an absolute constant. Particularly, when $r = n$ (i.e., X is sparse and l1-column-flat) then

$$W^2(\Gamma_X; \Phi_{A,B}) \leq 1 + n \log^2(n).$$

Theorem 3 Suppose $A_{ki}\sim iid B_{ij}\sim iid N(0,1)$ and independent each other, $X$ is a column-wise $s$-sparse and l1-column-flat signal, $Y = AX^B + E$ with measurement errors bounded by $|E|_F^2 \leq \eta$, $X^*$ is the minimizer of the problem $MP(\beta)_{Y,A,B,\eta}$. If

$$m \geq t + 4\sqrt{2}n^r \eta + C_1(ns)^{1/2} \log(C_2n^r).$$

where $C_i$’s are absolute constants, then $P[|X^* - X|_F \leq \delta] \geq 1 - \exp(-t^2/2)$, i.e., X can be reconstructed robustly with respect to the error norm $|X^* - X|_F$ with high probability by solving $MP(\beta)_{Y,A,B,\eta}$.

4.2. Case 2: A and B are both sub-Gaussian
This is the most complicated case. The main results are presented in the following. Foundations are in [9,10] and the complete proof is in the full version [11].

Lemma 4 Given $n$-by-$n$ matrix $X=(x_1,\ldots,x_n)$ with $s$-sparse column vectors $x_1,\ldots,x_n$, $r$ is cardinality of $\{j: |x_j|_1=\max_i|x_i|_1\}$, $\Phi_{A,B}, \Gamma_X, W(\Gamma_X; \Phi_{A,B})$ are specified as before, $A_{ki}\sim iid$ Sub-Gaussian distribution and $B_{ij}\sim iid$ Sub-Gaussian distribution with $\psi_2$-norms $\sigma_A, \sigma_B$ respectively, then

$$W^2(\Gamma_X; \Phi_{A,B}) \leq \sigma_A^2 \sigma_B^2 (1 + n^2 - r(n - \log^2(Cn^r))).$$

where $C$ is an absolute constant. Particularly, when $r = n$ then

$$W^2(\Gamma_X; \Phi_{A,B}) \leq \sigma_A^2 \sigma_B^2 (1 + n \log^2(Cn^3)).$$

Theorem 4 Suppose random matrices $A, B$ are independent each other, $A_{ki}\sim iid$ Sub-Gaussian distribution, $B_{ij}\sim iid$ Sub-Gaussian distribution, each with $\psi_2$-norm $\sigma_A$ and $\sigma_B$. Let $X$ be a columnwise $s$-sparse and l1-column-flat signal, $Y = AX^B + E$ with $|E|_F^2 \leq \eta, X^*$ be the minimizer of $MP(\beta)_{Y,A,B,\eta}$. If

$$m \geq t + 4\sqrt{2}n^r \eta + C_1 (\sigma_A \sigma_B)^{1/2} \log(C_2n^r)$$

where $C_i$’s are absolute constants, then $P[|X^* - X|_F \leq \delta] \geq 1 - \exp(-t^2/2)$, i.e., matrix signal X can be reconstructed robustly with respect to the error norm $|X^* - X|_F$ with high probability by solving $MP(\beta)_{Y,A,B,\eta}$.

Remark This result can be potentially applied to the important problem of subspace confinement.

5. Summary
As one of the most critical ingredients in dealing with the problem of robustly reconstructing structured matrix signal via convex programming in random setting, this paper established some fundamental lower bounds on number of random measurements to guarantee high successful reconstruction probability. In future works, these results will be generalized to more challenging and more general problem of tensor signal reconstruction, which is emerging as one of the critical problems in big data analysis, machine learning and many other related fields.

References
[1] S.Foucart, H.Rauhut, A Mathematical Introduction to Compressive Sensing, Birkhaeuser, 2013.
[2] Y.C.Eldar, G.Kutynoik, Compressed Sensing, Cambridge University Press, 2012.
[3] D.Cohen, Y.C.Eldar, Sub-nyquist radar systems: temporal, spectral and spatial compression, IEEE Signal Processing Magazine, 11(2018) 35-57.
[4] M.A.Davenport, J.Romberg, An overview of low-rank matrix recovery from incomplete observations, arXiv: 1601.06422, 2016.
[5] M.F.Duarte, R.G.Baraniuk. Kronecker compressive sensing, IEEE Transactions on Image Processing, 2012, 21(2):494-504.

[6] J.A.Tropp, Practical sketching algorithms for low-rank matrix approximation, SIAM J. Matrix Anal. Appl., 38(2017) 1454-1485.

[7] V.Chandrasekaran, B.Recht, P.A.Parrio, A.S.Wilsky, The convex geometry of linear inverse problems, Found. Comput. Math. 12(2012) 805-849.

[8] J.A.Tropp. Convex recovery of a structured signal from independent random linear measurements, in:G.Pfander(Ed.), Sampling Theory: A Renaissance, Birkhaeuser, 2015.

[9] S.Mendelson, Learning without concentration. J. ACM., 62(2014) 556-587.

[10] S.Mendelson, A.Pajor, N.Tomczak-Jaegermann, reconstruction and subgaussian operators in asymptotic geometric analysis, Geom. Func. Analysis, 27(2017) 1248-1282.

[11] Y. Tian, X. Huang: Convex reconstruction of structured matrix signals from linear measurements: theoretical results, preprint, arXiv:1910-08771, 2019.