IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received November 7, 2019, accepted November 19, 2019, date of publication December 4, 2019,
date of current version December 18, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2957479

A New Approach to Design Sensing Matrix Based on the Sparsity Constant With Applications to Computed Tomography

HAMID SHAHBAZI[1], MOHAMMAD REZA TABAN[1,2], (Senior Member, IEEE), AND JAMSHID ABOUEI[1], (Senior Member, IEEE)

[1]Department of Electrical Engineering, Yazd University, Yazd 89158-18411, Iran
[2]Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan 84156-83111, Iran

Corresponding author: Mohammad Reza Taban (mrtaban@cc.iut.ac.ir)

ABSTRACT Using random sensing matrices imposes some constraints on applying compressed sensing in practical applications such as computed tomography, high resolution radars, synthetic aperture radars and other imaging systems. On the other hand, the lack of certain criteria to measure the suitability of a sensing matrix in compressed sensing, makes designing of the relevant sampling system difficult; so, researchers have turned largely toward trial and error methods for designing such sensing matrices. In this paper, we propose a constructive approach to design measurement matrices which largely overcomes the aforementioned drawbacks and presents some simple and calculable measures for sensing matrices. The presented algorithm outperforms various random and deterministic approaches in designing compressed sensing matrices due to its recoverability performance and generality, at the same time, our scheme benefits from the fact that the sensing matrix performance is easily determined. Based on the proposed method and despite all constrains that exist in measurement matrices in different problems, we can design sensing matrices in a unified manner with even better performance than that of random scenarios. In addition, we apply the proposed algorithm in computed tomography where the measurement matrix is a structured matrix and our method can gain much improvement in the recoverability performance. Furthermore, this article provides parameters that affect the performance of a sensing matrix, which theoretically clarifies the causes of the good recoverability performance of Gaussian matrices and the poor recoverability performance of periodic matrices.

INDEX TERMS Compressed sensing, sparsity constant, measurement matrix design, transitivity of correlations, computed tomography.

I. INTRODUCTION

Compressed sensing (CS) has been widely adopted in imaging systems due to its high resource management efficiency and superior performance in sparse scenarios [1], [2]. CS combines two phases of data acquisition and data compression of signal $\mathbf{x} \in \mathbb{R}^n$ in a linear projection $\mathbf{y} = \mathbf{A}\mathbf{x}$ where $\mathbf{A}$ denotes the measurement or sensing matrix and $\mathbf{y}$ is the measured data. The sampling rate can be reduced in this linear projection which is important in various scenarios such as computed tomography (CT) and other medical imaging systems. For such applications, one can benefit some of the following advantages through employing CS theory: *i*) Reduction of the time needed for the measurement, *ii*) Saving the energy of sampling or reduction of the side effects of this energy on human body such as exposure to x-ray in CT, *iii*) Reduction of the required high speed memory, *iv*) Reduction of the required bandwidth for transferring sampled data, and *v*) Simplifying the analog to digital converter. On the other hand, when a number of samples is missed for different reasons, the CS theory can be used to recover the original signal without any degradation. Despite all of these advantages, there are some challenges in deploying the CS theory in practical applications such as CT and high resolution radars that motivate us to go further for a better design of CS-based measurement matrices.

The associate editor coordinating the review of this manuscript and approving it for publication was Pietro Savazzi.

A. RELATED WORKS

Many literature have been devoted to the usage of compressed sensing in different applications including computed tomography [2]–[6], high resolution radars [7], synthetic aperture radars [8], [9], and ground penetration radars [8]. In one line of these researches, the performance of different measurement matrices is often considered for various areas, such as high resolution radars [7], magnetic resonance imaging (MRI) [10], and ultra-wideband (UWB) channel estimation [11]. In another perspective, reference [9] compares the performance of different types of measurement matrix **A** including periodic structures, pure random structures and jitter random structures in synthetic aperture radars. Candès and Tao in [12] use the restricted isometry property (RIP) to formulate the recoverability of the CS problem. Based on the RIP analysis, in another line of relevant researches, a measurement matrix structure is introduced and its performance is examined by considering its RIP-ness or mutual coherence property with statistical or deterministic tools [13]–[15]. From the RIP analysis aspect, there are three schemes to design the sensing matrix[1] **A**:

i. **Random matrix approaches**: It is shown in [16] that the RIP-ness in random matrix approaches is hold with a high probability for different distributions, e.g., Gaussian matrices.

ii. **Deterministic methods**: In these approaches, the sensing matrix **A** with a good RIP-ness is designed by various deterministic methods such as number theory [17], approximation theory [18], expander graph theory [19], and coding theory [20]. A different approach for designing a binary sensing matrix based on the average of columns cross correlations is presented in [21].

iii. **Hybrid methods**: The sensing matrix **A** is assumed to be deterministic in these methods, while signal **x** is modeled as a random vector. For instance, reference [22] defines the statistical restricted isometry property (StRIP) constant instead of the RIP constant and then designs the sensing matrix **A** using the StRIP in a deterministic way.

B. MOTIVATION

Designing the sensing matrix **A** using deterministic approaches is of great interest due to the following reasons:

1) In almost all practical scenarios such as computed tomography, radars and MRI, the physics of the problem determines the measurement matrix structure [7], [10]. Thus, we cannot use random matrices in such applications.
2) The implementation of a random matrix is complex in some applications such as radars [22].
3) For random matrices with small dimensions, fine tuning of the measurement matrix is essential [23].
4) Random matrices may not display a proper performance in some applications where the quality of service is needed to be guaranteed [14].

[1]Throughout the paper, we occasionally use the term sensing matrix **A** instead of measurement matrix **A**.

In practical applications, the main challenge of applying, the compressed sensing theory is to design its measurement matrix and evaluate its performance from the recoverability points of view. The CS theory presents the use of random matrices and provides criteria such as the RIP constant and the null space property [24], however, it is shown that the calculation of these parameters is computationally intractable [25].

Taking the above considerations into account, the accomplishment of proposed methods for designing the measurement matrix should be investigated. To the best of our knowledge, in the most relevant research studies, the RIP-ness of the measurement matrix has been proved asymptotically and there is no proposed matrix that can beat random Gaussian matrices from the recoverability performance points of view. Some drawbacks of the RIP analysis will be presented in Section II. On the other hand as far as we know, the existence of a measurable criterion for evaluating the performance of the measurement matrix is an open problem that has not yet been solved. Furthermore, the construction of deterministic matrices with a better performance than random matrices is important in our understanding of the CS theory.

C. CONTRIBUTION

In this paper, we define the sparsity constant of the sensing matrix and use a constructive approach to design the sensing matrix **A** in a deterministic manner. Contrary to the computational intractability of the RIP constant and the null space property, the sparsity constant of the sensing matrix **A** can be easily calculated and used. In this method, we construct the measurement matrix **A** iteratively to achieve the maximum sparsity constant. The designed sensing matrix not only displays a good recoverability performance, similar to random matrices, but also in different scenarios, it outperforms the random matrix in terms of the recoverability performance. In addition, the proposed scheme answers an old question in the CS theory about the very poor performance of periodic sampling scenarios and the unified good recoverability of all random matrices. Furthermore, we employ the mean restricted isometry property (MRIP) introduced in Section IV to consider the robustness of the designed measurement matrix against noise. The proposed algorithm can be applied in almost every application that compressed sensing is used. Despite this generality, we also apply our algorithm to computed tomography in Section III and as we will show in this section, the proposed method can gain much improvement over recoverability performance. Numerical results verify the effectiveness of our approach and further confirm the validity of our theoretical analysis.

The rest of the paper is organized as follows. In Section II, we briefly describe some basic concepts of the CS theory and review its precise notations. The proposed sparsity constant based measurement matrix design is given in Section III. In Section IV, we deploy the proposed method to computed tomography. The conclusions are presented in Section V.

TABLE 1. Notations.

$x(k)$	The k^{th} elements of vector $\mathbf{x}$
$\|\mathbf{x}\|_{\ell_p}$	The $l_p norm$ of $\mathbf{x}$, defined as $(\sum_k x(k)^p)^{(1/p)}$
$\|\mathbf{A}\|_{\ell_p}$	the matrix norm defined as the according vector norm on all the elements of matrix $\mathbf{A}$.
$\mathbf{A}_{mn}$	The matrix $\mathbf{A}$ with m rows and n columns
$\mathbf{X}^T$	The transpose of matrix $\mathbf{X}$
$S_{\mathbf{A}}(k)$	The submatrix of matrix $\mathbf{A}$ with k columns
$\|S\|$	the cardinality of set S
$supp(\mathbf{x})$	the set of indices of non-zero elements of $\mathbf{x}$
$\mathbb{N}(\mathbf{A})$	The null space of matrix $\mathbf{A}$ defined as $\{\mathbf{x} \in R^n \| \mathbf{Ax} = \mathbf{0}\}$
$\mathbb{R}^n$	the n-dimensional Euclidean space
$gram(\mathbf{X})$	the Gram matrix of matrix $\mathbf{A}$ define as $Gram(\mathbf{A}) = \mathbf{A}^T\mathbf{A}$
$\lambda_min(\mathbf{A})$	the smallest eigenvalue of matrix $\mathbf{A}$
$E[f]$	The expectation operator of random variable f
$sign(x_i)$	The sign of x_i defined as $sign(x_i) = \begin{cases} +1 & x_i > 0 \\ 0 & x_i = 0 \\ -1 & x_i < 0 \end{cases}$
$N(0, \sigma^2)$	The normal distribution with variance σ^2

II. INTRODUCTION TO COMPRESSED SENSING

Compressed sensing asserts that we can measure sparse signals with fewer samples than what the Shannon sampling theorem requires [26]. CS relies strongly on the sparsity of measured signals meaning that if a signal is sparse, one can find a proper basis such that the sparse signal has a concise representation. Sampling in compressed sensing means a linear projection of a digital signal that typically has lower dimensions. In real scenarios, signals are analog and CS can be employed to sample signals just like the usual way and find a proper basis such that the data has a sparse representation based on it. Then, the CS process reduces dimensions of the signal using some linear projections.

A. NOTATIONS

In this paper, vectors and matrices are shown in bold lowercase and uppercase letters, respectively, and sets are shown by non-bold mathematical blackboard uppercase letters. The other notations is grouped in Table 1.

Let $\mathbf{x_0} \in \mathbb{R}^n$ represents a discrete k-sparse signal (i.e. $\|\mathbf{x}\|_{\ell_0} \leqslant k$), which must be measured and the measured data vector $\mathbf{b} \in \mathbb{R}^m$ is a linear projection of $\mathbf{x_0}$ sampled by the rows of sensing matrix $\mathbf{A}$. Compressed sensing in its original form uses the basis pursuit [26] to recover the original signal $\mathbf{x_0}$; thus, we can formulate the measurement and the recovery phases of compressed sensing as

$$\mathbf{A}\mathbf{x}_0 = \mathbf{b}, \tag{1}$$

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell 1} \quad \text{vs} \quad \mathbf{Ax} = \mathbf{b}. \tag{2}$$

Furthermore, $\mathrm{CR} = \frac{m}{n}$ is the compression ratio, and for each CS problem with sensing matrix $\mathbf{A}$ and k-sparse signals, triple (m, n, k) is defined as the problem size vector. A CS problem is recoverable if for a sensing matrix $\mathbf{A}$ and a problem size (m, n, k) , $\mathbf{x}_0 = \hat{\mathbf{x}}$ can be achieved from (2) with a high probability.

It is shown in [27] that the recoverability of a CS problem for a specific $\mathbf{x_0}$ depends only on the support and signs of $\mathbf{x_0}$. Thus, for a k-sparse signal, there are $\binom{n}{k} 2^k$ different signs and supports that among them only N number of them are recoverable. Hence, the probability of recoverability would be as

$$\mathrm{P}_{re}(k) = \frac{N}{\binom{n}{k} 2^k}. \tag{3}$$

In this paper, we frequently plot this measure versus k to compare the recoverability performance of different sensing matrices.

B. DIFFERENT TOOLS IN DEVELOPING CS THEORY

1) Approximation and measure theory: These approaches had been introduced based on the approximation theory in [28] and then have been considered again in the Donoho's work in [29]. As a brief explanation, let $\tilde{\mathbb{S}}$ denotes the set of all m-dimensional subspaces of $\mathbb{R}^n$ where CS is recoverable for sufficient sparse signals. It can be shown that the measure of $\tilde{\mathbb{S}}$ approaches to 1 when $(n - m) \to \infty$ (see [18] and [23] for more details). According to this theory, if we choose sensing matrix $\mathbf{A}$ in a uniform random form, then for sufficiently large size problems, the selected matrix will be a proper recoverable matrix with a high probability.

2) RIP analysis: The RIP constant of matrix $\mathbf{A}$ with s degree is the smallest number δ_s such that

$$(1 - \delta_s) \|\mathbf{x}\|_{\ell_2}^2 \leqslant \|\mathbf{Ax}\|_{\ell_2}^2 \leqslant (1+\delta_s) \|\mathbf{x}\|_{\ell_2}^2, \tag{4}$$

is held for every arbitrary s-sparse vector $\mathbf{x}$. It is shown that if $\delta_{2s} \geq \sqrt{2} - 1$, then matrix $\mathbf{A}$ is recoverable for all s-sparse vectors [24].

3) Null space property: Some drawbacks of the RIP analysis and the necessary and sufficient conditions for recoverability of CS with sensing matrix $\mathbf{A}$ have been given in [24] which depend on the null space of matrix $\mathbf{A}$. To clarify the subject, we just quote the following Theorem in [24] without any proof.

Theorem 1: Given $\mathbf{A} \in \mathbb{R}^{m \times n}$ and any integer $k \geqslant 1$, the recoverability

$$\{\mathbf{x}_0\} = \arg\min_{\mathbf{x}} \{\|\mathbf{x}\|_{\ell_1} : \mathbf{Ax} = \mathbf{A}\mathbf{x}_0\}, \tag{5}$$

holds for all $\mathbf{x}_0 \in \mathbb{R}^n$ such that $\|\mathbf{x}_0\|_{\ell_0} = k$ if and only if

$$\forall\ \mathbf{v} \neq \mathbf{0} : \mathbf{Av} = \mathbf{0} \Longrightarrow \|\mathbf{v}\|_{\ell_1} > 2\, \|\mathbf{v}_\Gamma\|_{\ell_1} \tag{6}$$

holds for all index sets $\Gamma \subset \{1, \cdots, n\}$ such that

$$|\Gamma| = k, \tag{7}$$

where $\mathbf{v}_\Gamma$ is the same as $\mathbf{v}$ for indexes included in set Γ and zero in other elements. Theorem 1 demonstrates that the recoverability of the compressed sensing problem depends only on the null space of matrix $\mathbf{A}$. If two matrices have the same null space, they have the same recoverability as well.

For example, row operations of a matrix do not change its null space; thus, we can use row operations to change the matrix to a reduced row echelon form of the matrix without any changing in the recoverability.

III. SPARSITY CONSTANT BASED MEASUREMENT MATRIX DESIGN

In this section, we propose a new method called sparsity constant based measurement (SCBM) matrix design that obtains the sensing matrix $\mathbf{A}$ iteratively based on its null space. This method provides a unified manner to design sensing matrices with different structures. Another advantage of this method is that the performance of a new designed sensing matrix can be measured easily. For such a scheme, we face with a subspace selection problem for the null space of matrix $\mathbf{A}$. Each point in space $\mathbb{R}^n$ is spanned by orthogonal bases $\mathbf{a}_1, \ldots, \mathbf{a}_n$. Accordingly, designing of matrix $\mathbf{A}$ means the selection of $n - m$ bases as the null space of matrix $\mathbf{A}$, and simply constructing matrix $\mathbf{A}$ with taking other m bases as the row vectors of matrix $\mathbf{A}$.

In most situations, the matrix structure is imposed by the problem mechanism. For instance, in a frequency spectrum estimation problem, the rows of matrix $\mathbf{A}$ may be the Fourier coefficients for determination of the Fourier transform and in a radar application, the columns of matrix $\mathbf{A}$ may be the shifted chirp signals. Taking the above considerations into account, the main contributions of this work are summarized as follows:

1) The recoverability of measurement matrices is managed by the sparsity constant (SP) in a sense that recoverability is improved with an increase in SP. Suppose that $\mathbf{B}$ is a matrix whose columns are orthonormal and span the null space of the measurement matrix $\mathbf{A}$, then the SP of $\mathbf{A}$ is defined as follow:

$$\mathrm{SP}(\mathbf{A}) = \|\mathbf{B}\mathbf{B}^{\mathbf{T}}\|_{\ell_1}. \tag{8}$$

2) SP($\mathbf{A}$) considers just the null space of matrix $\mathbf{A}$ and matrices designed based on sparsity constant might be too noise sensitive, although this condition occurs rarely. Thus, for designing a robust measurement matrix against noise in addition to sparsity constant, we must consider MRIP. Suppose the measurement matrix $\mathbf{A}$ with the problem size $(m,\ n,\ k)$ is recoverable based on its null space. Under this condition, the distribution of all eigenvalues of k-columns sub-matrices of $\mathbf{A}$ determines its performance against the noise. Nonetheless for most applications, the average of the smallest eigenvalue of k-columns sub-matrices of sensing matrix $\mathbf{A}$, is a good criterion to express the matrix robustness against the noise. Thus, we propose the MRIP of a measurement matrix $\mathbf{A}$ with k-sparse signals as follows:

$$\mathrm{MRIP}(\mathbf{A}, k) = \mathrm{E}[\lambda_{\min}\,(\mathrm{S}_{\mathbf{A}}(k))\,], \tag{9}$$

where E denotes the expectation operator and $S_{\mathbf{A}}(k)$ is a k-columns sub-matrix of sensing matrix $\mathbf{A}$. In general, we should continue the simulations until the observed variations of the estimated quantity is stabilized, i.e., it tends to a steady state situation. In practice, the averaging of $\lambda_{\min}\,(\mathrm{S}_{\mathbf{A}}(k))$ over few hundreds randomly selected sub-matrices is sufficient for estimating the reliable MRIP.

A. SOME IMPORTANT DISCUSSIONS

Before going to why we propose sparsity constant, we should discuss some important questions for clearance:

1) Although constructing sensing matrices is one of the heavily researched topics, but there still exist much works to do. Note that the measure theory and the RIP analysis are two different methods which come from different phenomena that are used to prove the same problem in the asymptotic case $(n - m) \to \infty$. In practice, when CS uses a sensing matrix in a special problem size, the dominant factor is usually unknown. As an example, one can refer to distributions of the smallest and largest eigenvalues of Gaussian sensing matrices given in [30] and [31]. Using these distributions, we can consider the RIP constant for a Gaussian matrix. Under this condition and similar to the results represented in [12], the RIP analysis guarantees that the CS problem is recoverable with the Gaussian matrix for an 5-sparse signal and $\frac{1}{2}$ compression ratio, if the signal length is more than 14000. Such a RIP analysis originates from some phenomena that are essentially different from the case when we use CS to recover 30-sparse signals with the signal length 200. Of course, these method are very rigorous but are very restrictive, underestimate the recoverability, computationally intractable and do not give any insight to design measurement matrices.
2) Due to the aforementioned drawbacks, the mutual coherence of the sensing matrix, $\mu(\mathbf{A})$ proposed for measuring of the recoverability. Let $\mathbf{G} = \hat{\mathbf{A}}^T\hat{\mathbf{A}}$, where $\hat{\mathbf{A}}$ is the column normalized version of $\mathbf{A}$, then $\mu(\mathbf{A})$ can be stated as:

$$\mu(\mathbf{A}) = \max_{i \neq j} \mathbf{G}(i,j). \tag{10}$$

It is shown that all k-sparse vector is recoverable when k satisfies the following inequality [26]:

$$k < 1/2(1 + 1/\mu(\mathbf{A})). \tag{11}$$

For efficient optimizing of $\mu(\mathbf{A})$, [32] proposes minimizing of

$$\overline{\mu(\mathbf{A})} = \|\mathbf{G} - \mathbf{I}\|_{\ell_2}, \tag{12}$$

where $\mathbf{I}$ is identity matrix. Since $\mathbf{G}$ is the gram of the column normalized matrix, $\mu(\mathbf{A})$ and $\overline{\mu(\mathbf{A})}$ are two parameters against each other and usually a reduction in $\overline{\mu(\mathbf{A})}$ leads to an increase in $\mu(\mathbf{A})$ for a well designed sensing matrix. The idea of [32] is a misunderstanding and as far as we examined, low coherency for compressed sensing is a useless scheme. Indeed, the sparsity

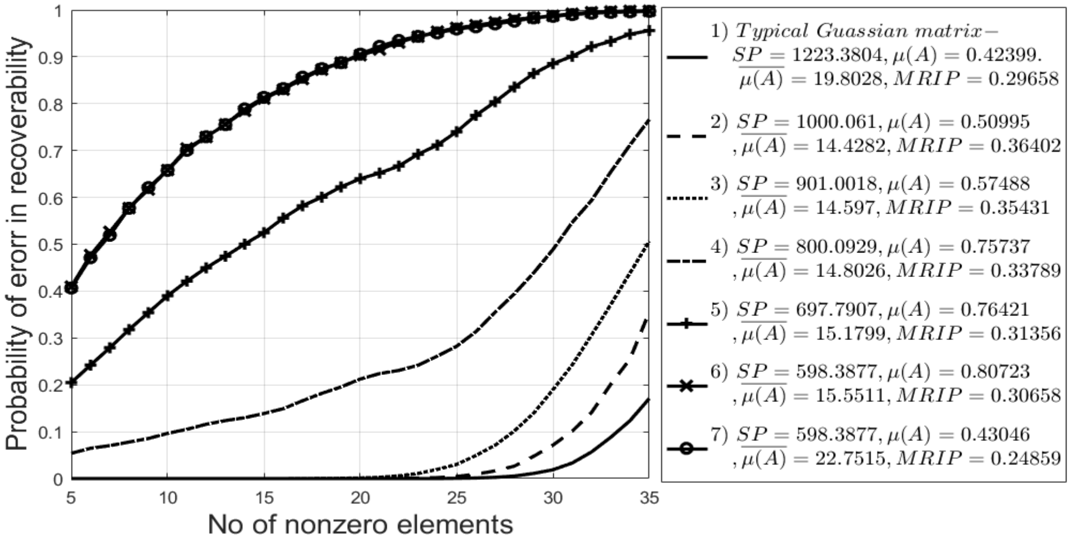


FIGURE 1. The recoverability performance of Gaussian manipulated matrices with $m = 100$, $n = 200$ for different sparsity constants. As it is shown, the SP constant is the major parameter in recoverability.

constant is neither the low coherency scheme nor its dual for the following reasons:

i) For a typical Gaussian measurement matrix with $m = 5000$ and $n = 10000$, the inequality (11) states that just 5-sparse signal is recoverable while all 1900-sparse signal can be recovered efficiently, thus, inequality (11) does not even state anything about the real recoverability of Gaussian random matrices.

ii) We manipulate randomly a Gaussian matrix so that we decrease its sparsity constant while keeping $\mu(\mathbf{A})$ and $\overline{\mu(\mathbf{A})}$ unchanged. The recoverability performance error is shown in Fig. 1. As seen in this figure, recoverability is decreased with reduction of the sparsity constant while $\mu(\mathbf{A})$ and $\overline{\mu(\mathbf{A})}$ have not effect on it.

3) Designed matrix with a high sparsity constant may be noise sensitive, thus, MRIP constant must take into consideration not meaning that the MRIP constant is sufficient. MRIP in our theory is just an axillary parameter. MRIP is similar to the RIP constant which is based on the infimum of λ_k, the minimum eigenvalue of k column sub matrices of $\mathbf{A}$, and MRIP is based on the average of all λ_k, thus, MRIP unlike RIP is computationally tractable. The major parameter is sparsity constant that is one can design matrices with high MRIP constant, low sparsity constant and poor recoverability as it is shown in Fig. 1. However in the optimization and designing the measurement matrix, we must consider that the MRIP does not decrease so that the designed matrices with the proposed algorithm gain improvement in recoverability performance without any reduction in noise robustness.
4) The gained improvement in recoverability performance with the proposed method is seen for both noiseless and noisy scenarios as it is shown in Fig. 2. It should be noted that, in Fig. 2 the recoverability phase is based on the following semidefinite program using the second order cone programming [33].

Measurement phase:

$$\mathbf{b} = \mathbf{A}\mathbf{x}_0 + z, \qquad z \sim N(0, \sigma^2) \tag{13}$$

Recovery phase:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell 1} \quad \text{vs} \quad \|\mathbf{A}\mathbf{x} - \mathbf{b}\| \leqslant \alpha E[\|\mathbf{z}\|_{\ell_2}], \tag{14}$$

where α is a regulation parameter.
5) Sometimes the original signal is not sparse, but there is a dictionary $\boldsymbol{\phi}$ that signal $\mathbf{x}$ has a sparse representation in it. Then, the overall measurement matrix would be as follows:

$$\left.\begin{aligned} \boldsymbol{A}\mathbf{x} &= \boldsymbol{b} \\ \mathbf{x} &= \boldsymbol{\phi} z \end{aligned}\right\} \Longrightarrow \boldsymbol{A}\boldsymbol{\phi}\mathbf{x} = \boldsymbol{b}. \tag{15}$$

It should be noted that the sparsity constant is applicable for each case where the full rank, total measurement matrix $\mathbf{A}_{mn}$ is ill, i.e., $m < n$. Thus, when signal $\mathbf{x}$ has an over complete sparsity basis $\boldsymbol{\phi}_{np}$ that is $p > n$, the total measurement matrix, $\boldsymbol{A\phi}$ is ill and the sparsity constant is applicable either with orthogonal and non-orthogonal dictionary.
6) The MRIP and SP are computationally tractable, nicely fitted recoverability performance, giving us efficient tools in destining compressed sensing but do not foresee the recoverability region. However, comparison of these parameters with the according parameters of the same size Gaussian matrix, give us handy tools.

B. PRELIMINARIES

In this subsection, we consider the sparsity property, ℓ_0-norm, ℓ_1-norm and their relations to sparsity. When a vector $\mathbf{x}$ is

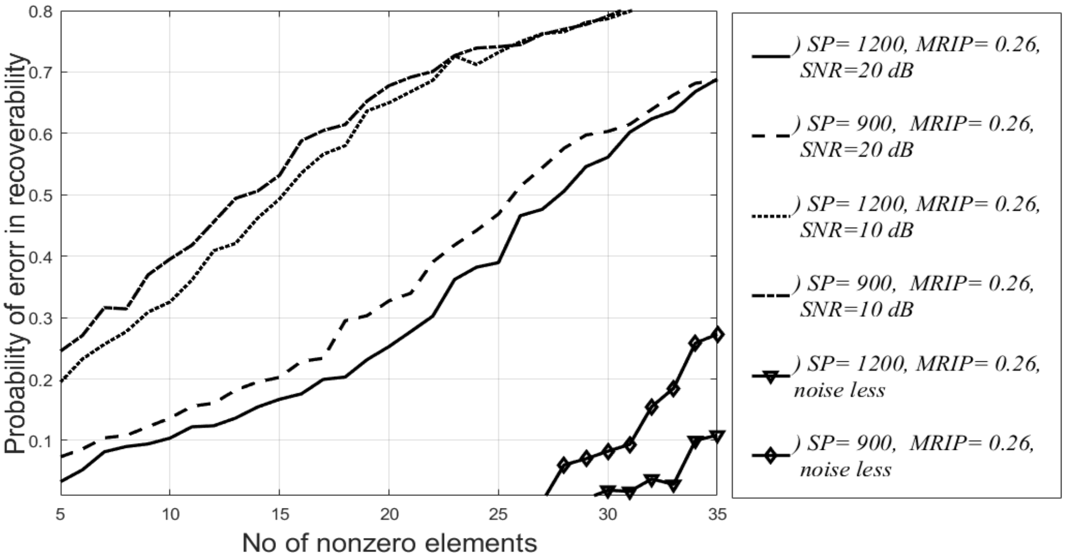


FIGURE 2. The recoverability performance of Gaussian manipulated matrices with $m = 100$, $n = 200$ for different sparsity constants and various SNRs. As could be seen from the figure, the proposed method improve recoverability performance in noisy and noiseless scenarios.

sparse, it means that the major part of its energy is concentrated on some few of its elements. Thus, when a vector is multiplied by a constant coefficient, it does not change the sparsity behavior of the vector. For this reason, we suppose that all vectors are normalized to ℓ_2-norm, from the sparsity points of view. For normalized vectors, ℓ_1-norm is an excellent measure for investigating the sparsity of a vector; i.e., less ℓ_1-norm of a vector implies more sparsity of the vector and vice versa. The maximum ℓ_1-norm of a normalized vector occurs when all its elements have the same magnitude, while the minimum occurs when its whole energy is concentrated on a single element.

Consider two normalized vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ with the same elements except the first two elements as follows:

$$\mathbf{v}_1 = [\mathrm{a}_1, \mathrm{a}_2, \cdots, \mathrm{a}_\mathrm{n}], \tag{16}$$

$$\mathbf{v}_2 = [\mathrm{a}_1 + \alpha, \mathrm{a}_2 + \beta, \cdots, \mathrm{a}_\mathrm{n}]. \tag{17}$$

Without loss of generality, let α and a_1 have the same signs, while, β and a_2 have the opposite signs, $|\mathrm{a}_1| \leqslant |\mathrm{a}_2|$ and $\max(|\alpha|, |\beta|) \ll \big||\mathrm{a}_1| - |\mathrm{a}_2|\big|$. In order to keep $\mathbf{v}_2$ normalized, α and β cannot simultaneously have equal (or opposite) signs with a_1 and a_2 and it is easy to show that $|\alpha| \geqslant |\beta|$, which means that $\|\mathbf{v}_2\|_{\ell_1} \geqslant \|\mathbf{v}_1\|_{\ell_1}$. Thus for a normalized vector, reducing its larger element and increasing its smaller one, i.e., a_2 and a_1, lead to an increase in its ℓ_1-norm. Therefore, ℓ_1-norm is an excellent measure for sparsity, if its ℓ_2-norm is constant. In fact, increasing ℓ_1-norm while keeping ℓ_2-norm constantly, distributes the energy between different elements of the normalized vector. In this paper, we use this property to define the sparsity constant and design the sensing matrix $\mathbf{A}$ based on its null space. On the other hand, this property can use for gaining the sparse solution of an undetermined equation system shown as follows:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1}, \tag{18}$$

$$\mathbf{Ax} = \mathbf{Ax_0}, \tag{19}$$

$$\|\mathbf{x}\|_{\ell_2} = \|\mathbf{x_0}\|_{\ell_2}. \tag{20}$$

This set of equations is non-convex and indissoluble. Nonetheless, we can ignore (20) because $\mathbf{Ax} = \mathbf{b}$ plus the RIP-ness property of measurement matrices, approximately provide the ℓ_2-norm constraint.

C. SPARSITY CONSTANT

Let consider a k-sparse vector denoted by $\mathbf{x}$. According to Theorem 1, to satisfy the non-recoverability of $\mathbf{x}$, there must exist a vector $\boldsymbol{\alpha}$ in the null space of $\mathbf{A}$ such that its elements with opposite signs with some of the corresponding k nonzero elements of $\mathbf{x}$ follow

$$\mathbb{M} = \{i : |x_i| > 0, \quad \mathrm{sign}(x_i\alpha_i) = -1\}, \tag{21}$$

$$\sum_{i \in \mathbb{M}} |\alpha_\mathrm{i}| \geqslant \sum_{i \in \mathbb{M}^c} |\alpha_i|, \tag{22}$$

where $\mathbb{M}^c$ denotes the complement of set $\mathbb{M}$. In contrast, if there exists a vector $\boldsymbol{\alpha}$ with the property (22), then all sparse vectors which have the opposite sign elements with vector $\boldsymbol{\alpha}$ in indexes included in set $\mathbb{M}$, are not recoverable by the ℓ_1 programming.

Definition 1: A point in the null space that satisfies the property (22) is defined as the ill point with degree $|\mathbb{M}|$. The mean of a point is a specific support of the vector and its signs. Thus, we just need to consider the supports and signs of a vector in $\mathbb{R}^n$.

Lemma 1: Let assume

$$\min_i (|\alpha_{i \in \mathbb{M}}|) \geqslant \max_i (|\alpha_{i \in \mathbb{M}^c}|). \tag{23}$$

Then, the inequality in (24) is a necessary condition for vector $\boldsymbol{\alpha}$ to be an ill point in the index of set $\mathbb{M}$.

$$\sum_{i\in\mathbb{M}}\alpha_i^2 \geqslant \sum_{i\in\mathbb{M}^C}\alpha_i^2. \tag{24}$$

Proof: See Appendix A.

This lemma shows that the accumulation of the energy in a few number of elements makes an ill point. Thus, if we avoid the accumulation of energy in the design of the null space of a measurement matrix, we can reduce the number of ill points. It can be shown that an ill point with degree $|\mathbb{M}|$ makes $2^{k-|\mathbb{M}|}\binom{n}{k-|\mathbb{M}|}$ points from all $2^k\binom{n}{k}$ k-sparse points, irrecoverable. The number of ill points is a fundamental property of a subspace meaning that it does not depend on any specific basis of the subspace.

Let $\mathbb{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ denotes a set of orthonormal bases of $\mathbb{R}^n$. If $n - m$ vectors of them can span the proper null space[2] of the desired measurement matrix, then all points of the null space can be shown as

$$\begin{aligned}&\mathbf{y} = [\mathbf{a}_1, \cdots, \mathbf{a}_{n-m}]\,\mathbf{x},\\ &\mathbf{x}\in\mathbb{R}^{n-m} \text{ and } \|\mathbf{x}\|_{\ell_2} = 1,\end{aligned} \tag{25}$$

where $\mathbf{y}$ spans the null space of the measurement matrix. Since, we only need to consider normalized vectors and all the basis vectors are orthonormal, we drop $\|\mathbf{x}\|_{\ell_2} = 1$ in (25). Let us rewrite (25) in the row direction as follows:

$$\begin{bmatrix} y_1\\ \vdots \\ y_n\end{bmatrix} = \begin{bmatrix}\mathbf{b}_1^T\\ \vdots\\ \mathbf{b}_n^T\end{bmatrix}\mathbf{x}. \tag{26}$$

If $[y_1, \cdots, y_n]^T$ is an ill point with indexes of set $\mathbb{M}$, then there is a vector $\mathbf{x}$ that has significant correlations with $\{\mathbf{b}_J |\, J\in\mathbb{M}\}$ and lower correlations with other vectors. Hence, the following necessary and sufficient conditions can be deduced for good recoverability performance of the measurement matrix.

Necessary condition: For reducing the number of ill points, it is necessary that every group of vectors $\mathbf{b}_J, J \in \mathbb{M}$, has a good cross correlation with all other group of vectors. If there is a group of vectors $\mathbf{b}_{j_1}^T, \ldots, \mathbf{b}_{j_s}^T$ which are orthogonal or near orthogonal to another group of vectors like $\mathbf{b}_{j_{s+1}}^T, \ldots, \mathbf{b}_{j_n}^T$, then, we can choose $\mathbf{x}$ in the span of $\mathbf{b}_{j_1}^T, \ldots, \mathbf{b}_{j_s}^T$, and there will be ill points in the null space. Thus, the non-orthogonally of some groups of $\mathbf{b}_1^T, \ldots, \mathbf{b}_n^T$ to others is an essential condition for the recoverability of the subspace. In another way, we can use the Van der Corput Lemma [34] for showing the property of transitivity of correlation, which basically asserts that if a vector $\mathbf{x}$ is correlated with many vectors of $\mathbf{b}_1^T, \ldots, \mathbf{b}_n^T$, then, many pairs of $\mathbf{b}_1^T, \ldots, \mathbf{b}_n^T$ are correlated with each other.

Sufficiency of good cross correlations of $\mathbf{b}_1^T, \ldots, \mathbf{b}_n^T$ for good recoverability: Increasing the cross correlations between vectors $\mathbf{b}_1^T, \ldots, \mathbf{b}_n^T$, is sufficient for decreasing the number of ill points. Suppose that $\mathbf{x}$ has a good correlation with $\{\mathbf{b}_{j_1}^T\}$ and very poor correlations with $\{\mathbf{b}_{j_2}^T, \ldots, \mathbf{b}_{j_s}^T\}$, and these two subsets have a good correlation with each other. The vector $\mathbf{x}$ can be indicated as a sum of two orthogonal vectors $\mathbf{b}_{j_1}$ and $\mathbf{b}_{j_1}^{\perp}$ as

$$\mathbf{x} = \alpha\mathbf{b}_{j_1} + \alpha^c\mathbf{b}_{j_1}^{\perp}, \tag{27}$$

where α and α^c represent the correlation between $\mathbf{x}$ and two vectors $\mathbf{b}_{j_1}$ and $\mathbf{b}_{j_1}^{\perp}$, respectively. Then, it is sensibly expected that with an increase in the cross correlations between $\mathbf{b}_{j_1}$ and $\mathbf{b}_{j_2}, \ldots, \mathbf{b}_{j_s}$, the cross correlations between $\mathbf{x}$ and $\mathbf{b}_{j_2}, \ldots, \mathbf{b}_{j_s}$ increase, which reduces the number of ill points.

Based on the above arguments, for reducing the number of ill points in the null space, the cross correlations between all pairs of $\mathbf{b}_{j_1}, \ldots, \mathbf{b}_{j_s}$ should increase as much as possible. Thus, to analyze the recoverability of the sensing matrix $\mathbf{A}$, we should study the cross correlations between the rows of null space matrix $\mathbf{b}_{j_1}^T, \ldots, \mathbf{b}_{j_n}^T$. These cross correlations can be shown with the Gram matrix[3] $\mathbf{G}$ as follows

$$\begin{aligned}\mathbf{AB} &= \mathbf{0},\\ \mathbf{B}^T\mathbf{B} &= \mathbf{I},\\ \mathbf{G} &= \mathbf{BB}^T,\end{aligned} \tag{28}$$

where $\mathbf{A}$ is the $m \times n$ sensing matrix, $\mathbf{B}$ is a $n \times (n - m)$ matrix with orthogonal columns which its columns span the null space of matrix $\mathbf{A}$ and $\mathbf{I}$ represents the identity matrix. The interesting property of the Gram matrix $\mathbf{G}$ is that every criterion for the recoverability of sensing matrix $\mathbf{A}$ should depend only on the null space of matrix $\mathbf{A}$. It should be noted that $\mathbf{G}$ in (28) depends on the null space of matrix $\mathbf{A}$ and does not change with the row operations on matrix $\mathbf{A}$. On the other hand, ℓ_2-norm of $\mathbf{G}$ is equal to ℓ_2-norm of rows $\mathbf{b}_1^T, \ldots, \mathbf{b}_n^T$, which is constant and equals to $m - n$. Thus, for increasing the cross correlations between these vectors, ℓ_1-norm of $\mathbf{G}$ should increase.

D. SPARSITY CONSTANT BASED MEASUREMENT MATRIX DESIGN ALGORITHM

Taking the above considerations into account, one can use the following problem for designing measurement matrices with optimized recoverability. Let an orthogonal basis $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$, spans the space $\mathbb{R}^n$. The optimized subspace with k dimensions (as null space of measurement matrix $\mathbf{A}$), is a set of vectors $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_k}\}$, which its Gram matrix has the maximum ℓ_1-norm, i.e,

$$\text{Optimized null space} = \underset{i_1,\cdots,i_k}{\arg\max}\left\{\left\|\text{Gram}\left(\left[\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_k}\right]\right)\right\|_{\ell_1}\right\}, \tag{29}$$

where $\text{Gram}(\mathbf{A}) = \mathbf{A}^T\mathbf{A}$. This is a very complicated combinatorial optimization problem which is not solvable. However, we can use greedy algorithms to find an approximate solution.

[2] A proper null space is a null space with a reduced number of ill points.

[3] Some of the Gram matrix characteristics is given in [35].

Greedy constructive measurement matrix design based on sparsity constant: Using greedy approaches, one can construct sensing matrix $\mathbf{A}$ as follow:

i) Choose an initial $\mathbf{a}_{i_1}$ from $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ and set $k = 1$.
ii) $\mathbb{S} = \{i_1\}$ and $\hat{\mathbb{S}} = \{i_2, \ldots, i_n\}$.
iii) $k \leftarrow k + 1$.
iv) $i_k = \arg\max_{i_k \in \hat{\mathbb{S}}} \left\{ \left\| \mathrm{Gram}\left[\left[\mathbf{a}_{\{i|i \in \mathbb{S}\}}\right], \mathbf{a}_{i_k}\right] \right\|_{\ell_1} \right\}$.
v) $\mathbb{S} = \mathbb{S} \cup \{i_k\}$, $\hat{\mathbb{S}} = \hat{\mathbb{S}} - \{i_k\}$.
vi) Repeat step *ii* to step *v*.
vii) Stop when k=n-m.

It should be noted when there are some constraints such as when the elements of matrix must be binary or when the non-compressed matrix is structured, the greedy optimization problem changes the elements of it in a manner that the constraint is satisfied in each step; e.g. when the there is a structured non-compressed matrix, the greedy algorithm change the selected rows of the non-compressed matrix in each step (e.g., for the chirp matrix in Fig. 3) and when the elements of matrix must be binary the greedy algorithm, just changes the sign of one of the elements in the measurement matrix in each step. We denote $\|\mathbf{G}\|_{\ell_1}$ of the null space of measurement matrix $\mathbf{A}$ as the sparsity constant of the matrix represented by SP$(\mathbf{A})$. We use this parameter to compare the recoverability of different measurement matrices in compressive sensing. The recoverability performance error $P_{re}(k)$ of different structures and different problem sizes of a measurement matrix $\mathbf{A}$ is shown in Fig. 3. In this figure, CR is the compression ratio, DCT is the discrete cosine transform matrix and k is the number of non-zero elements. In our simulation, we define a bad matrix as the random matrix with the lowest sparsity constant in 10000 iterations, good matrix is a random matrix with the highest sparsity constant in 10000 iterations, and "OPT" matrix is the matrix that is optimized by the proposed algorithm. The value after each matrix name (e.g., Bad, 1190 in Fig. 3) is the sparsity constant of the matrix. It is worth mentioning that the sparsity constant of a bad ($m = 100, n = 200$) DCT matrix is 1168, while the sparsity constant of a same size periodic compressed DCT matrix is 460. This very low sparsity constant justifies the poor recoverability performance of the periodic compressed measurement matrices.

In Fig. 4, similar to Fig. 3, the proposed method is applied to the binary, trinary and the Gallager LDPC code by simulations. As be seen, the proposed method can be used successfully and effectively either for considering the recoverability performance and also optimizing the desired measurement matrix. In the LDPC case, the improvement is very little because the constraints in the columns and rows of the measurement matrix decrease the degree of freedom during the optimization of the sparsity constant.

E. ROBUSTNESS OF SPARSITY CONSTANT BASED MEASUREMENT MATRIX DESIGN

A detailed discussion about the performance analysis of the CS theory against the noise has been given in [36], [37]. Based on the null space property, we analyze the effect of an additive noise for the proposed method. The noise changes measurement vector $\mathbf{b} = \mathbf{A}\mathbf{x}_0$ to $\mathbf{b}' = \mathbf{A}\mathbf{x}_0 + \mathbf{z}$, where $\mathbf{z} \sim N(0, \sigma^2)$ represents the additive noise vector. In this regard, we have the following three scenarios:

1) CS without noise: This is the original noiseless CS problem, i.e.,

$$\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1} \quad st. \quad \mathbf{A}\mathbf{x} = \mathbf{b}. \tag{30}$$

2) Robust CS (RCS): When the noise exists, the constraint equation $\mathbf{A}\mathbf{x} = \mathbf{b}$ cannot display the real constraint equations, even these set of equations may not be consistent; thus, we must substitute the constraint $\mathbf{A}\mathbf{x} = \mathbf{b}$ with $\|\mathbf{A}\mathbf{x} - \mathbf{b}'\|_{\ell_2} \leqslant \varepsilon$, i.e.,

$$\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1} \quad st. \quad \|\mathbf{A}\mathbf{x} - \mathbf{b}'\|_{\ell_2} \leqslant \varepsilon. \tag{31}$$

3) Non-robust CS (NRCS) when noise exists: In this mode, we use the original CS recoverability scheme, despite of the noise existence, i.e.,

$$\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1} \quad \text{st.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}'. \tag{32}$$

If measurement matrix $\mathbf{A}$ has a good RIP property, it can be shown that the solution is approximately correct [36]. It is worth mentioning that with a non-RIP but a good null space measurement matrix[4], the recoverability performance of the NRCS decreases drastically as can be seen in Fig. 5.

In the first scenario, CS is recoverable if vector $\mathbf{x}$ is the sparsest vector in the affine space $\mathbf{A}\mathbf{x} = \mathbf{b}$. For recoverability of the robust CS, $\mathbf{x}_0$ must be the sparsest vector in a family of closed affine subspaces $\|\mathbf{A}\mathbf{x} - \mathbf{b}\| \leqslant \varepsilon$. In this case, the error can occur in two ways, *i)* vector $\mathbf{x}$ is not a member of these subspaces, *ii)* another sparser vector can be found in these subspaces.

As mentioned earlier, the recoverability in the noiseless scenario depends only on the null space of the measurement matrix. However, this problem is completely different from noisy situations. In this case, the robustness against the noise is reduced by decreasing the minimum eigenvalue of sub-matrices composed by a number of the sensing matrix columns. Let ε denotes the amplitude of the added noise. If ε is in the direction of the minimum absolute eigenvalue λ_{min}, then the noisy term, which can be added to $\mathbf{x}$, has $\frac{\varepsilon}{\lambda_{\min}}$ amplitude. This problem makes subspaces $\|\mathbf{A}\mathbf{x} - \mathbf{b}'\|_{\ell_2} \leqslant \varepsilon$ too large and consequently the recoverability will reduce drastically. Therefore, a second measure namely, MRIP constant, is required for the recoverability of a matrix when the noise exists.

Mean restricted isometry property: Suppose the measurement matrix $\mathbf{A}$ with the problem size (m, n, k) is recoverable based on its null space. Under this condition, the distribution of all eigenvalues of k-columns sub-matrices of $\mathbf{A}$

[4] A good null space measurement matrix is a matrix with a high sparsity constant which implies a good recoverability performance.

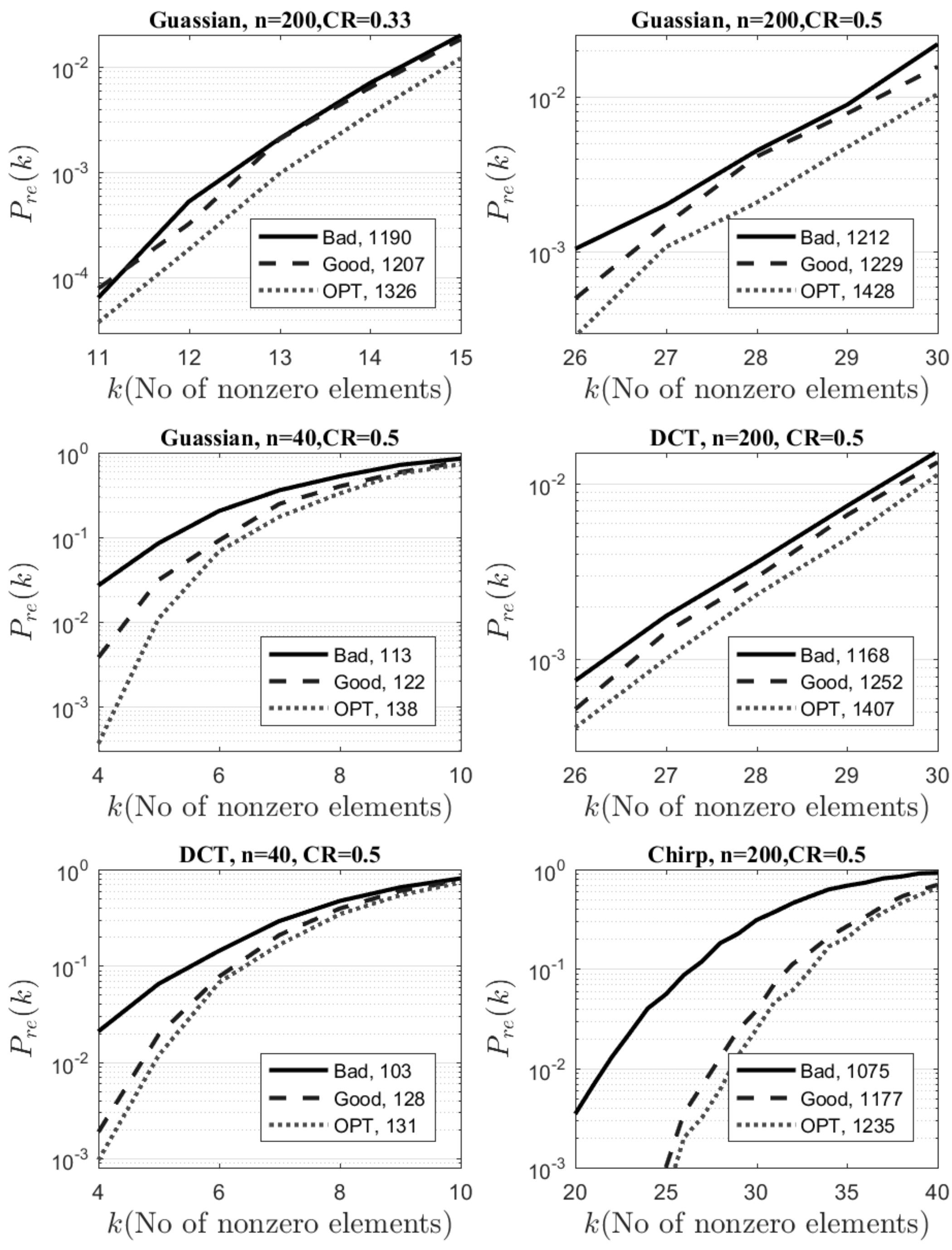


FIGURE 3. **The recoverability performance error of different structures of a measurement matrix A with various problem sizes. The curves have been prepared for three cases of A: bad matrix, good matrix and optimized (proposed) matrix. The sparsity constant of each case is presented in the legend.**

determines its performance against the noise. Nonetheless for most applications, the average of the smallest eigenvalue of k-columns sub-matrices of sensing matrix $\mathbf{A}$, is a good criterion to express the matrix robustness against the noise. Thus, we propose the MRIP of a measurement matrix $\mathbf{A}$ with k-sparse signals as has been defined in (9). Unlike the RIP constant problem which has been proved to be NP-hard [25], the MRIP constant is computationally tractable. For example, using the random optimization method and starting from a typical Gaussian matrix, the sparsity constant of the matrix can be optimized. Although the performance of this matrix in the noiseless scenario is better than that of random matrices, its performance reduces drastically when the noise exists. Therefore, to design a noise resistant matrix, we use the

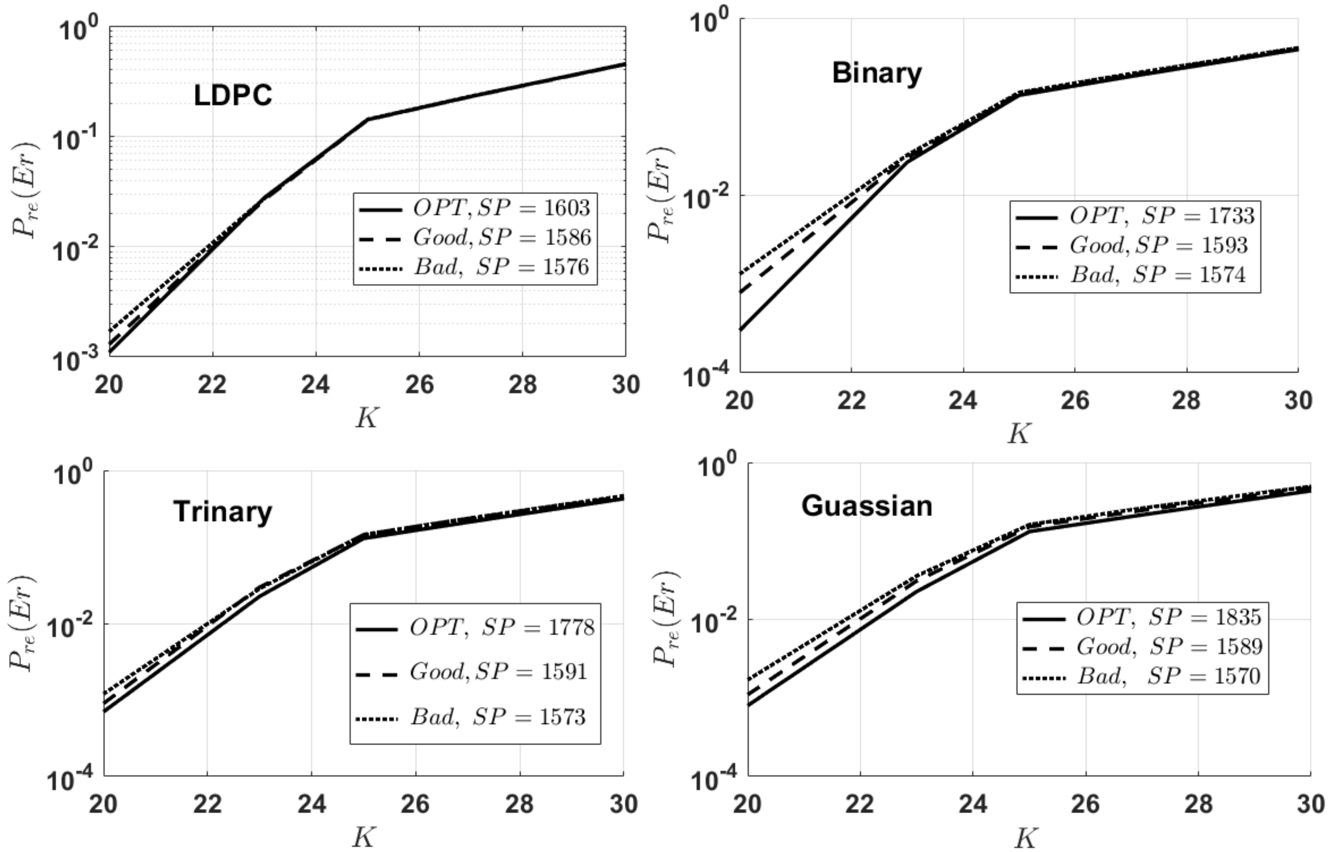


FIGURE 4. The recoverability performance error of a measurement matrix with different coding structures including Binary, Trinary and Gallager LDPC codes in comparison to a Gaussian random matrix with similar size. The curves have been prepared for three cases of bad matrix, good matrix and optimized (proposed) matrix. The sparsity constant of each case is presented in the legend.

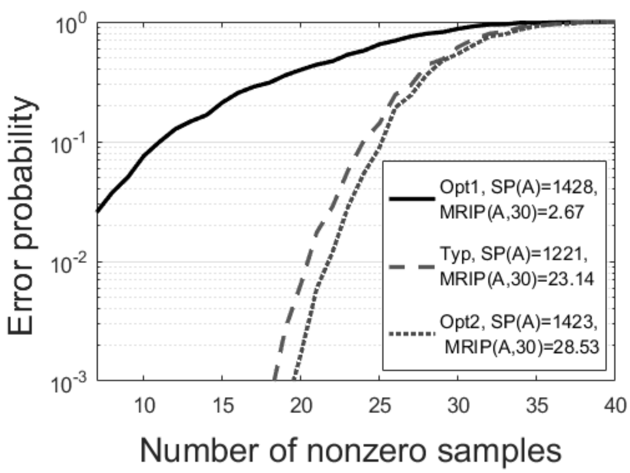


FIGURE 5. The recoverability performance of a Gaussian CS matrix with SNR = 20dB and CR = 0.5 from different methods: typical random (Typ), optimized without the MIRP consideration (Opt1), and optimized when the MIRP constant is held (Opt2).

same optimization method to maximize the sparsity constant of the matrix, while, we must avoid decreasing the MRIP constant. This optimization problem is numerically evaluated in Fig. 5 for the Gaussian matrix, SNR = 20 dB and the compression ratio CR = 0.5. In this figure, "Typ" is a typical

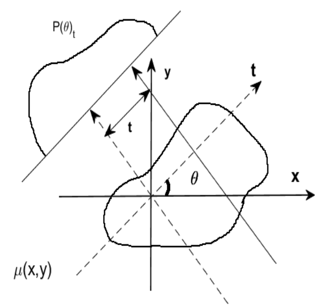


FIGURE 6. Parallel beam geometry tomography setup [38].

random matrix, "opt1" is an optimized sparsity constant matrix without the MRIP consideration and "opt2" is an optimum matrix with considering the MRIP constant. For all cases, signals are recovered by the RCS using the second order cone programming [33].

F. DISCUSSION ON THE IMPLEMENTATION ISSUES

There three different measurement matrix schemes in compressed sensing: i) deterministic matrices, ii) quasi random matrixes and iii) true random matrices. In most situations, the used measurement matrices are quasi random or deterministic where the proposed algorithm can be

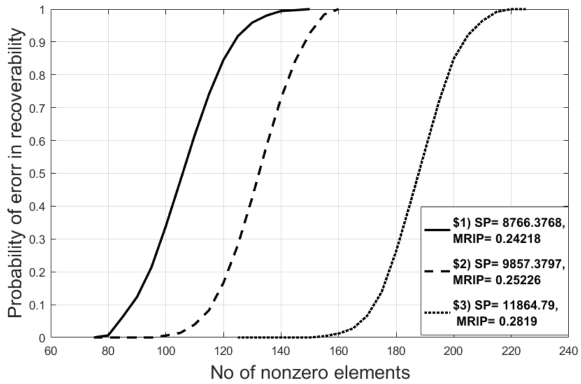


FIGURE 7. Recoverability performance error of 1) bad, 2) good and 3) opt matrices versus the number of nonzero elements. As it is shown the proposed method has been able to increase the recoverable region by 77%.

used effectively. In these situations, greedy algorithms are used to design measurement matrices offline. Then, with this offline designed measurement matrix, we can implement measurement simply while we are sure that our measurement setup has a good recoverability performance. Note that a true random measurement matrix is not a good technical choice, since one cannot predict the recoverability of the system beforehand. Although for some cases (e.g., Gaussian random matrices), the variance of recoverability performance is small, however, in many situations such as chirp matrix in radar application or Radon transform matrix in computed tomography, the variance of recoverability performance is large. In a real random measuring system, the measurement matrix is not designed beforehand and the measurement matrix is revealed to us simultaneous to the end time of the measurement, then we use this matrix with one of the different recovering algorithms such as basis pursuit to recover original signal. Finally, we would like to emphasize to this point that even in true random matrices, we can use sparsity constant effectively to evaluate the recoverability performance of the measurement matrix simply.

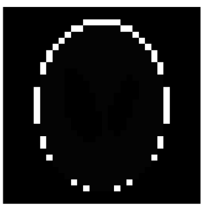
IV. COMPRESSED COMPUTED TOMOGRAPHY: CASE STUDY

Computed tomography (CT) provides tools for investigating the inner structure of an object and has many medical and industrial applications [38]. The idea behind CT is to find the image of the 2 dimensional attenuation function, $\mu(x, y)$ of an object from its projection in different angles. The setup of the 2D parallel beam geometry tomography is shown in Fig. 6.

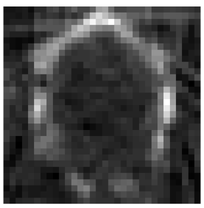
It can be shown that the logarithm of attenuation function is $\mathcal{R}[\mu(x, y)]$, the Radon transform of $\mu(x, y)$ [38]. Thus, as it is shown in Fig. 6, in parallel beam computed tomography one can obtain $P_\theta(t)$, the projection of $\mu(x, y)$ in the angle θ and position t as follow:

$$P_\theta(t) = \mathcal{R}_\theta[\mu(x, y)] = \iint_{-\infty}^{\infty} \mu(x, y)\,\delta(xcos\theta + ysin\theta - t)dxdy, \qquad (33)$$

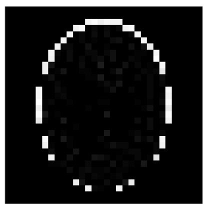
where $\delta(t)$ is the Dirac function. Using the CS for finding $\mu(x, y)$ from $P_\theta(t)$ when $\mu(x, y)$ is sparse, requires to



(a) Phantom Shepp-Logan.

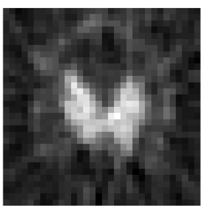


(b) Shepp-Logan, conjugate gradient algorithm with Tikhonov regularization

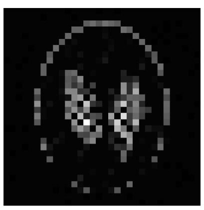


(c) Shepp-Logan, the recovered image with the bad matrix.

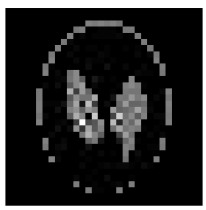
(d) Less sparse image.



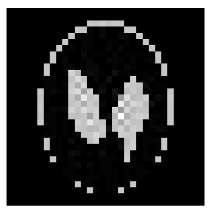
(e) Less sparse image, conjugate gradient algorithm with Tikhonov regularization



(f) Less sparse image, the recovered image with the bad matrix.

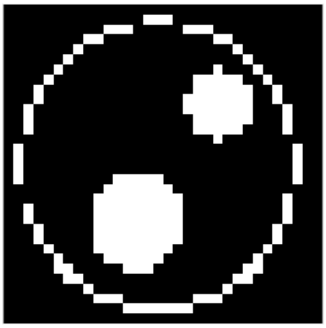


(g) Less sparse image, the recovered image with the good matrix.

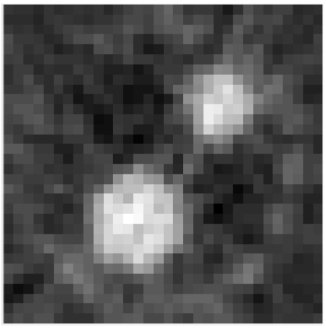


(h) Less sparse image, the recovered image with the opt matrix.

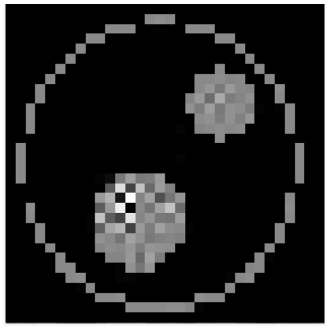
FIGURE 8. The recoverable image of Shepp-Logan and one artificiality less sparse image with different reconstruction methods. As it is shown, the compressed sensing recovering method outperforms significantly other methods with very few number of measurements when the scene is sparse and the proposed method reduces the sparsity requirements substantially.



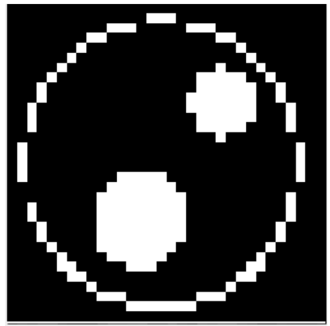
(a) two bubbles phantom.



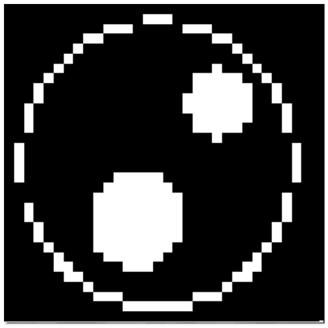
(b) Conjugate gradient algorithm with Tikhonov regularization.



(c) Compressed sensing with random matrix.



(d) Compressed sensing with good random matrix chosen with the proposed method.



(e) Compressed sensing with optimized sparsity constant (the proposed algorithm)

FIGURE 9. The recoverable image of two bubbles phantom image [41] with different reconstruction methods.

discretize (33) to a linear equation form:

$$\mathbf{Az} = \mathbf{s}, \tag{34}$$

where $\mathbf{A}$,$\mathbf{z}$ and $\mathbf{s}$ will be defined shortly. Suppose that $\mu\,(x, y)$ is discretized with $x_i = x_1 + i\Delta x \;\; i = 1, \ldots, N$, and $y_i = y_1 + i\Delta y \;\; i = 1, \ldots, N$, then the matrix $\boldsymbol{\mu}$ is defined as follow:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu(x_1, y_1) & \cdots & \mu(x_1, y_N) \\ \vdots & \ddots & \vdots \\ \mu(x_N, y_1) & \cdots & \mu(x_N, y_N) \end{bmatrix}.$$

$\mathbf{z}$ is the column base vectorized version of the $\boldsymbol{\mu}$ matrix or in the Matlab language: $\mathbf{z} = \boldsymbol{\mu}\,(:)$. Let discretize the measurement angles and t-axis as follow:

$$\begin{aligned} \theta_i &= \Delta\theta \times i, \quad i = 1, \ldots, m, \\ t_j &= t_1 + \Delta t \times j, \quad j = 1, ..., n. \end{aligned} \tag{35}$$

Since, the Radon transform is linear, thus the Radon transform of a scene is the sum of Radon transform of each pixel [39]. For any measurement in the angle θ, each column of $\mathbf{A}_\theta$ is the discretized Radon transform when only the according pixel intensity is equal to one and the others are zero. Thus, $\mathbf{A}_{\theta_i}$ and $\mathbf{z}_{\theta_i}$ are defined as follow:

$$\mathbf{A}_{\theta_i} = \Big[R_{\theta_i}[\delta\,(x - x_1, y - y_1)], \ldots\; R_{\theta_i}[\delta\,(x - x_i, y - y_i)] \;\ldots, R_{\theta_i}[\delta\,(x - x_N, y - y_N)] \Big], \tag{36}$$

$$s_{\theta_i} = P_{\theta_i}(t_1, ..., t_n), \tag{37}$$

and the final measurement matrix become:

$$\underbrace{\begin{bmatrix} A_{\theta_1} \\ \vdots \\ A_{\theta_N} \end{bmatrix}}_{\mathbf{A}} [\mathbf{z}] = \underbrace{\begin{bmatrix} s_{\theta_1} \\ \vdots \\ s_{\theta_N} \end{bmatrix}}_{\mathbf{s}}. \tag{38}$$

For applying the CS theory, we must omit more rows of (38), which means the reduction in the measurements number. Since we do not know in prior any information of the object, the easiest way is the random selection of some rows of $\mathbf{A}$. Without SCBM, one cannot even consider the recoverability performance of a selected random matrix. We make measurement matrix for an (32 × 32) image with just 500 measurements, with 10000 random iterations, in which *i*) choose the matrix with the most high sparsity constant and it is named as "good", *ii*) the most low sparsity constant and it is named as "bad", *iii*) then we make one compressed measurement matrix with the proposed greedy algorithm with the name "opt". The recoverability error of these matrices is shown in Fig. 7. As seen from this figure, there exists a large improvement in the recoverability performance in these matrices with the increase in the sparsity constant.

Now we use these different matrices in recovering the head phantom Shepp-Logan Test image (Fig. 8a)) which has been used widely by researchers and it is a standard test image [40] and one artificial less sparse image (Fig. 8d)). The result is shown in Fig. 8. As it can be seen from Fig. 8, even the bad matrix can recover 32×32 resolution of Shepp-Logan Test image, while the popular conjugate gradient algorithm with Tikhonov regularized method cannot reconstruct the image with these very few measurements. However, bad and good matrices cannot recover an artificial less sparse image in Fig. 8(d), but the opt matrix can recover it without degradation in Fig. 8(h).

Further, we consider image reconstructions for the two bubbles phantom image similar to [41] in the parallel beam computed tomography in Fig. 9. The recoverable images are shown in sub-figures 9(b), (c), (d) and (e) using the following reconstruction methods, respectively: *i*) conjugate gradient algorithm with Tikhonov regularization *ii*) compressed sensing with random matrix. *iii*) compressed sensing with good random matrix chosen with the proposed method *iv*) compressed sensing with optimized sparsity constant (the proposed algorithm). As it is shown, the compressed sensing recovering method outperforms significantly other methods with very few number of measurements. As be seen, not every measurement matrix can be used properly for two bubbles phantom; however, the proposed method enables us to select good measurement matrices or design a matrix with optimized sparsity constant. Although either of matrices that is used in sub-figures 9(d) and (e) recover the two bubbles Phantom image completely, the optimized matrix of sub-figures 9(e) even can recover less sparse image.

For the other schemes that use compressed sensing such as applying sparsity in total variation of the image [42], frequency domain [5] and Coded-Aperture X-Ray Luminescence Tomography [2], our measurement matrix design method can be used to reduce the sparsity requirements.

The implementation of the proposed algorithm in the CT is straightforward. In this regard because of hardware limitations of the CT, we should select a subset of $\theta_i s$ in designing a sensing matrix. Note that the measurements can also be done in specific values of both θ_i and x_i. Then in the implementation, we measure the intensity function only in these $\theta_i s$. When x_i values is selectable, we must mask the parallel beams to prevent the x-ray beams in the non-selected x_i as described in [2]. Each selection of the parameters generates a different sensing matrix with different capabilities. The proposed algorithm can be used to select the optimum parameters.

V. CONCLUSION

In this paper, we defined the sparsity constant of the measurement matrix and analyzed its effect on the recoverability performance. We used this new parameter as the heart of the recoverability performance, to propose a novel constructive approach for designing the measurement matrix based on SC. It was demonstrated that using the proposed algorithm, it is possible to design measurement matrix in a deterministic way with a very good recoverability performance even better than Gaussian measurement matrices. Moreover, we considered the MRIP constant (as a calculable and easy to use criterion) along with the null space property for analyzing the recoverability performance to design the noise robust matrix. The described theory explained an old question about the good recoverability of random matrices and poor recoverability of periodic sampling. In addition, the proposed method was applied successfully to the computed tomography where it could make a large improvement in recoverability performance. Overall, our theory paves the way of using compressed sensing in different applications where random matrices cannot be used and in different scenarios can improve the recoverability performance and reduce the sparsity requirements.

APPENDIX A
PROOF OF LEMMA 1

To prove the lemma, let assume $\alpha_i > 0$ and let

$$|\mathbb{M}| = \mathrm{k} + 1, \ \left|\mathbb{M}^{\mathrm{C}}\right| = \mathrm{m}. \tag{39}$$

We just consider the infimum situation and normalize case where $\sum_{\mathrm{i}\in\mathbb{M}} \alpha_\mathrm{i} = \sum_{\mathrm{i}\in\mathbb{M}^\mathrm{C}} \alpha_\mathrm{i} = 1$. Without loss of generality, let us assume that α_i is ordered from the greatest to least as follow:

$$\underbrace{\alpha_0,\ \alpha_1, \cdots, \alpha_k}_{\text{Set } \mathbb{M}},\ \underbrace{\alpha_{k+1}, \cdots, \alpha_{k+m}}_{\text{Set } \mathbb{M}^{\mathbb{C}}}. \tag{40}$$

Consider the necessary condition (41) which must be proved.

$$\sum_{\mathrm{i}\in\mathbb{M}} \alpha_\mathrm{i}^2 \geqslant \sum_{\mathrm{i}\in\mathbb{M}^\mathrm{C}} \alpha_\mathrm{i}^2. \tag{41}$$

The minimum value of the left hand side is when the energy of indexes in set $\mathbb{M}$ distributes equally as much as possible and the maximum value of the right hand side is where the energy of indexes in set $\mathbb{M}$ concentrates in a few number of them, as much as possible. Let $\alpha_k = \beta$, then this extreme case can be shown as follow:

$$\underbrace{\underbrace{\frac{1-\beta}{k}, \cdots, \frac{1-\beta}{k}}_{k},\ \beta,}_{\text{Set } \mathbb{M}}\ \underbrace{\underbrace{\beta, \cdots, \beta}_{k'},\ (1-k'\beta),\ \underbrace{0, \cdots, 0}_{(m\text{-}k')},}_{\text{Set } \mathbb{M}^{\mathbb{C}}} \tag{42}$$

where $k' = fix(1/\beta)$ and the normalize condition implies that $0 \leqslant \beta \leqslant 1/(k+1)$. Now the necessary condition in (41) is simplified as

$$\beta^2 + \frac{(1-\beta)^2}{k} \geqslant k'\beta^2 + (1-k'\beta)^2. \tag{43}$$

Let $m = 1 - k'\beta$, then the right hand side of (43) is less than β as follows:

$$k'\beta^2 + (1-k'\beta)^2 = (1-m)\beta + m^2 \leqslant (1-m)\beta + m\beta = \beta \tag{44}$$

Then (43) is converted to

$$\beta^2 + \frac{(1-\beta)^2}{k} \geqslant \beta. \tag{45}$$

It is easy to see that (45) is equal to $\beta \leqslant 1/(k+1)$.□

REFERENCES

[1] S. Zhang, J. Wu, D. Chen, S. Li, B. Yu, and J. Qu, "Fast frequency-domain compressed sensing analysis for high-density super-resolution imaging using orthogonal matching pursuit," *IEEE Photon. J.*, vol. 11, no. 1, Feb. 2019, Art. no. 6900108.

[2] S. Tzoumas, D. Vernekohl, and L. Xing, "Coded-aperture compressed sensing X-ray luminescence tomography," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 8, pp. 1892–1895, Aug. 2018.

[3] F. Yang, D. Zhang, K. Huang, W. Shi, and X. Wang, "Scattering estimation for cone-beam CT using local measurement based on compressed sensing," *IEEE Trans. Nucl. Sci.*, vol. 65, no. 3, pp. 941–949, Mar. 2018.

[4] C. G. Graff and E. Y. Sidky, "Compressive sensing in medical imaging," *Appl. Opt.*, vol. 54, no. 8, pp. C23–C44, Nov. 2015.

[5] Z. Zhu, K. Wahid, P. Babyn, D. Cooper, I. Pratt, and Y. Carter, "Compressed sensing-based algorithm for sparse-view CT image reconstruction," *Comput. Math. Methods Med.*, vol. 55, no. 12, pp. 5695–5702, Jan. 2013.

[6] W. Hou and C. Zhang, "A compressed sensing approach to low-radiation CT reconstruction," in *Proc. 9th Int. Symp. Commun. Syst., Netw. Digit. Sign (CSNDSP)*, Jul. 2014, pp. 793–797.

[7] M. A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2275–2284, Feb. 2009.

[8] S.-H. Jung, Y.-S. Cho, R.-S. Park, J.-M. Kim, H.-K. Jung, and Y.-S. Chung, "High-resolution millimeter-wave ground-based SAR imaging via compressed sensing," *IEEE Trans. Magn.*, vol. 54, no. 3, Mar. 2018, Art. no. 9400504.

[9] V. M. Patel, G. R. Easley, D. M. Healy, Jr., and R. Chellappa, "Compressed synthetic aperture radar," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 244–254, Apr. 2010.

[10] K. H. Jin, D. Lee, and J. C. Ye, "A general framework for compressed sensing and parallel MRI using annihilating filter based low-rank Hankel matrix," *IEEE Trans. Comput. Imag.*, vol. 2, no. 4, pp. 480–495, Dec. 2016.

[11] T. Yaacoub, O. A. Dobre, R. Youssef, and E. Radoi, "Optimal selection of Fourier coefficients for compressed sensing-based UWB channel estimation," *IEEE Wireless Commun. Lett.*, vol. 6, no. 4, pp. 466–469, Aug. 2017.

[12] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[13] Q. Jiang, S. Li, H. Bai, R. C. de Lamare, and X. He, "Gradient-based algorithm for designing sensing matrix considering real mutual coherence for compressed sensing systems," *IET Signal Process.*, vol. 11, no. 4, pp. 356–363, Jun. 2017.

[14] D. Bryant, C. J. Colbourn, D. Horsley, and P. Ó Catháin, "Compressed sensing with combinatorial designs: Theory and simulations," *IEEE Trans. Inf. Theory*, vol. 63, no. 8, pp. 4850–4859, Aug. 2017.

[15] X.-J. Liu, S.-T. Xia, and F.-W. Fu, "Reconstruction guarantee analysis of basis pursuit for binary measurement matrices in compressed sensing," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 2922–2932, May 2017.

[16] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[17] P. Indyk, "Explicit constructions for compressed sensing of sparse signals," in *Proc. 19th Annu. ACM-SIAM Symp. Discrete Algorithms*, 2008, pp. 30–33.

[18] R. A. DeVore, "Deterministic constructions of compressed sensing matrices," *J. Complex.*, vol. 23, nos. 4–6, pp. 918–925, Aug. 2007.

[19] V. Guruswami, J. R. Lee, and A. Razborov, "Almost Euclidean subspaces of ℓ_1^N via expander codes," in *Proc. 19th Annu. ACM-SIAM Symp. Discrete Algorithms*, Jan. 2008, pp. 353–362.

[20] M. A. Iwen, "Simple deterministically constructible RIP matrices with sublinear Fourier sampling requirements," in *Proc. 43rd Annu. Conf. Inf. Sci. Syst.*, Mar. 2009, pp. 870–875.

[21] W. Lu, T. Dai, and S.-T. Xia, "Binary matrices for compressed sensing," *IEEE Trans. Signal Process.*, vol. 66, no. 1, pp. 77–85, Jan. 2018.

[22] R. Calderbank, S. Howard, and S. Jafarpour. (2009). *Deterministic Compressive Sensing With Groups of Random Variables*. [Online]. Available: http://www.dsp.ece.rice.edu/files/cs/strip-more.pdf

[23] T. L. N. Nguyen and Y. Shin, "Deterministic sensing matrices in compressive sensing: A survey," *Sci. World J.*, vol. 2013, Sep. 2013, Art. no. 192795.

[24] Y. Zhang, "Theory of compressive sensing via ℓ_1-minimization: A non-RIP analysis and extensions," *J. Oper. Res. Soc. China*, vol. 1, no. 1, pp. 79–105, Mar. 2013.

[25] A. M. Tillmann and M. E. Pfetsch, "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1248–1259, Feb. 2014.

[26] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[27] D. L. Donoho and J. Tanner, "Counting faces of randomly projected polytopes when the projection radically lowers dimension," *J. Amer. Math. Soc.*, vol. 22, no. 1, pp. 1–53, Jul. 2008.

[28] B. S. Kashin, "Diameters of certain finite-dimensional sets in classes of smooth functions," *J. Amer. Math. Soc.*, vol. 41, no. 1, pp. 1–53, 1977.

[29] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[30] I. M. Johnstone, "On the distribution of the largest eigenvalue in principal components analysis," *Ann. Statist.*, vol. 29, no. 2, pp. 295–327, 2001.

[31] A. Edelman, "The distribution and moments of the smallest eigenvalue of a random matrix of wishart type," *Linear Algebra Appl.*, vol. 159, no. 3, pp. 55–80, Dec. 1991.

[32] H. Hu, M. Soltanalian, P. Stoica, and X. Zhu, "Locating the few: Sparsity-aware waveform design for active radar," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 651–662, Feb. 2017.

[33] K. C. Toh, M. J. Todd, and R. H. Tütüncü, "SDPT3—A MATLAB software package for semidefinite programming, version 1.3," *Optim. Methods Softw.*, vols. 11–12, nos. 1–4, pp. 545–581, Jan. 1999.

[34] T. Tao. (2014). *When is Correlation Transitive?* [Online]. Available: https://terrytao.word press.com/2014/06/05/when-iscorrelation-transitive/

[35] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. New York, NY, USA: Cambridge Univ. Press, 2012.

[36] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comp. Rendus Math.*, vol. 346, nos. 9–10, pp. 589–592, May 2008.

[37] E. J. Candés, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, May 2006.

[38] A. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging* (Classics in Applied Mathematics). Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2001. [Online]. Available: https://books.google.com/books?id=Z6RpVjb9_lwC

[39] J. Mueller and S. Siltanen, *Linear and Nonlinear Inverse Problems With Practical Applications*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2012. [Online]. Available: https://epubs.siam.org/doi/abs/10.1137/1.9781611972344

[40] L. A. Shepp and B. F. Logan, "The Fourier reconstruction of a head section," *IEEE Trans. Nucl. Sci.*, vol. NS-21, no. 3, pp. 21–43, Jun. 1974.

[41] K. Li, T. C. Chandrasekera, Y. Li, and D. J. Holland, "A non-linear reweighted total variation image reconstruction algorithm for electrical capacitance tomography," *IEEE Sensors J.*, vol. 18, no. 12, pp. 5049–5057, Jun. 2018.

[42] X. Han, J. Bian, D. R. Eaker, T. L. Kline, E. Y. Sidky, E. L. Ritman, and X. Pan, "Algorithm-enabled low-dose micro-CT imaging," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 606–620, Mar. 2011.

HAMID SHAHBAZI received the B.Sc. degree in communications engineering and the M.Sc. degree in communication systems engineering from Azad University, Iran, in 2008 and 2011, respectively. He is currently pursuing the Ph.D. degree in communication systems engineering with the Department of Electrical and Computer Engineering, Yazd University, Iran. He is also a full-time Researcher and Teaching Assistant with the Imaging Laboratory, Department of Electrical and Computer Engineering, Yazd University, Iran. His research interests are in general areas of signal processing and optimization, and their applications in various computed tomography and microwave imaging systems.

MOHAMMAD REZA TABAN was born in Isfahan, Iran, in 1968. He received the B.Sc. degree in electrical engineering from the Isfahan University of Technology (IUT), Isfahan, in 1991, the M.Sc. degree in electrical engineering from Tarbiat Modarres University, Tehran, Iran, in 1993, and the Ph.D. degree in electrical engineering from the Isfahan University of Technology, in 1998. He joined the Faculty of Yazd University, Yazd, Iran, in 1999. Since 2016, he has been a Professor of electrical and computer engineering at the Isfahan University of Technology, Isfahan. He has published more than 89 technical articles in international and national journals and conferences. His main research interests are statistical signal processing, detection, and estimation. He has been in the Scientific Committee of the 10th International ISC Conference (ISCISC2013) and the 2nd Iranian Conference on Avionics System (ICAS2015).

JAMSHID ABOUEI (SM'13) received the B.Sc. degree in electronics engineering and the M.Sc. degree (Hons.) in communication systems engineering from the Isfahan University of Technology (IUT), Iran, in 1993 and 1996, respectively, and the Ph.D. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 2009. From 1996 to 2004, he was a Faculty Member (Lecturer) with the Department of Electrical Engineering, Yazd University, and from 1998 to 2004, he was a Technical Advisor and Design Engineer (part-time) with the Research and Development Center, and the Cable Design Department, SGCC Company. From 2009 to 2010, he was a Postdoctoral Fellow with the Multimedia Laboratory, Department of Electrical and Computer Engineering, University of Toronto, ON, Canada. He is currently an Associate Professor with the Department of Electrical Engineering, Yazd University, Iran. His research interests are in the general areas of wireless ad hoc and sensor networks, with particular reference to energy efficiency and optimal resource allocation, multiuser information theory, and nanonetworks. He has received numerous awards and scholarships, including FOE and IGSA awards for excellence in research from the University of Waterloo, Canada, and the MSRT Ph.D. Scholarship from the Ministry of Science, Research and Technology, Iran, in 2004, and the Best Paper Award from the ICEE2018 Conference.