*Journal of AI and Data Mining*
***Vol 8, No 1, 2020, 13-23.***
DOI: 10.22044/JADM.2019.6908.1810

# A Sharp Sufficient Condition for Sparsity Pattern Recovery

Z. Shaeiri*, M.-R. Karami and A. Aghagolzadeh

*Department of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran.*

Received 18 March 2018; Revised 26 April 2018; Accepted 02 June 2019
**Corresponding author: z.shaeiri@stu.nit.ac.ir (Z. Shaeiri).*

**Abstract**
A sufficient number of linear and noisy measurements for exact and approximate sparsity pattern/support recovery in the high-dimensional setting is derived. Although this problem has been addressed in the recent literature, the results can still be improved. In this paper, a methodology is proposed, which reduces the gap between the existing upper bound on the error probability of support set recovery and the exact limit. This leads to a sharper sufficient condition for support recovery. A specific form of a joint typicality decoder is used for the support recovery task. Two performance metrics are considered for the recovery validation: one that considers the exact support recovery, and the other that seeks the partial support recovery. First, an upper bound is obtained on the error probability of the support recovery. Next, using the mentioned upper bound, a sufficient number of measurements for a reliable support recovery is derived. It is shown that the sufficient condition for a reliable support recovery depends on three key parameters of the problem: the noise variance, the minimum non-zero entry of the unknown sparse vector, and the sparsity level. Results are proved and examined both theoretically and by experiments. Simulations are performed for different sparsity rates, different noise variances. The results show that for all the mentioned cases, the proposed methodology increases the convergence rate of the error probability upper bound significantly.

**Keywords:** *Sparsity Pattern Recovery, Support Set Recovery, Information-Theoretic Limits, Performance Bound, Joint-Typicality Decoder, Compressed Sensing.*

## 1. Introduction

In signal processing, one commonly faces with an estimation problem in which a vector $x \in R^M$ must be estimated from a linear noisy observation vector denoted as:

$$y = Ax + n \tag{1}$$

where $A \in R^{N\times M}$ is known and $n \in R^N$ is the additive noise with a known distribution. It is known that $x$ can be estimated from $M$ measurements. However, when $x$ is sparse, it is possible to estimate it with a sufficient accuracy from a far fewer number of measurements, say $N \ll M$. This possibility is proved and further analyzed in compressed sensing [1-5]. $y$ is such output noisy measurement of the *k*-sparse vector $x$, i.e. it has $k < N \ll M$ non-zero entries, and $n$ is the additive noise vector. The support set of $x$ is defined as:

$$S = \{i \,|\, x_i \neq 0\} \tag{2}$$

In the problem of sparsity pattern or support set recovery, the concern is to detect non-zero positions of the vector $x$ [6] when we are given the measurement vector $y$ and the measurement matrix $A$. A reliable support recovery is important in many applications including magnetoencephalography (MEG), electroencephalography (EEG), cognitive radio, subset selection in regression, and multi-user communication systems [7-11]. The focus of many of the recent studies is on designing and analyzing the tractable recovery algorithms to solve (1) for $x$. Another brand of works studied the information-theoretic limits of any estimator for an exact or approximate recovery of the support of $x$ [12-24].

Information-theoretic limits disclose the extent of sub-optimality of the current sub-optimal methods. In other words, they reveal the gap between the solutions of the currently sub-optimal methods that obtain the sparsity pattern and the information-theoretic limits of the problem. Generally, the information-theoretic limits can be realized in two ways: necessary conditions and sufficient conditions, which give the support recovery conditions of two extreme cases. A lower bound on the number of measurements can be considered as the sufficient condition for an exact support recovery, i.e. if the number of measurements $N$ is more than $N_0$, the exact support recovery is guaranteed for an optimal decoder. On the other hand, an upper bound on the number of measurements can be viewed as a necessary condition for a support recovery, i.e. if the number of measurements $N$ reduces to $N_0$, then the exact support recovery is not possible using any decoder.

## 2. Relation to previous works

The problem of reconstruction or estimation of the sparse signal $x$ has attracted substantial attentions. Given the measurement model (1) and the sparsity assumption (2), one approach for estimating $x$ is to solve the $\ell_1-$ constrained quadratic program below, known as LASSO [31], given by:

$$\min_{x\in R^M}\{\frac{1}{2N}\|y-Ax\|_2^2+\lambda_N\|x\|_1\}$$

in which $\lambda_N>0$ is the regularization parameter. A great deal of recent works use $\ell_1-$ constrains for estimation of $x$ in the presence of sparsity constraints. In this direction of studies, several algorithms with affordable complexities have been proposed such as Orthogonal Matching Pursuit (OMP) [32], Subspace Pursuit (SP) [33], Compressive Sampling Matching Pursuit (CoSaMP) [34], and Iterative Hard Thresholding (IHT) [35].

In another brand of studies, only the sparsity pattern recovery is taken into account. In fact, in some applications, detecting the support set is an ultimate goal [7-11]. By determining the support set, the signal can be estimated simply by solving a least squares problem.

In this work, we are concerned with fundamental limits of the sparsity pattern recovery problem. Fundamental limits of a recovery problem can be achieved by analyzing the performance of an optimal decoder. These limits are highly valuable since they reveal the gap between the performance of any tractable recovery algorithm and the ultimate performance limits. In this work, a sufficient condition of support recovery is taken into account. In the recent years, in several works, the authors have studied the information-theoretic limits of any estimator for an exact and approximate support recovery and for single and multiple measurement vector models. Using an optimal decoder, Wainwright [19] has presented the necessary and sufficient conditions on $(k,N,M)$ for which an exact support recovery is possible in the high-dimensional setting. Akcakaya and Tarokh [20] studied the necessary and sufficient conditions for an exact support recovery in the high-dimensional setting using a certain type of joint typicality (JT) decoder for different error criteria. Under the assumption of finite constant per sample SNR (Signal-to-Noise Ratio) and measurement rate, Reeves and Gastpar [21] showed that an optimal recovery was possible with a constant fraction of error. Fletcher *et al.* [22] obtained the necessary condition for an exact support recovery using the maximum likelihood decoder in a certain setting. Xu *et al.* [23] developed the probability of a partial support recovery and asymptotic mean-square error of the recovered sparse signal for a maximum likelihood decoder. In [24], Rad derived sufficient conditions using the Chernoff technique and some features of the eigenvalues of the difference of the projection matrices.

In this work, the sufficient condition is enhanced. Most related to our work are the works of Akcakaya *et al.* [20] and Scarlett *et al.* [16]. Akcakaya proved that $O(k)$ measurements are sufficient for support recovery in the high-dimensional setting using a JT decoder. He proved that when $N\gg Ck$, where $C\geq 1$ was a constant, a perfect recovery was possible. Scarlett obtained sufficient conditions via the analysis of the JT decoder, while a prior distribution was assumed on the support set. In this work, we make the constant $C$ explicit. This constant is derived in terms of the key parameters of the problem. Actually, we derive a sharper sufficient condition for the sparsity pattern recovery when the same JT decoder is used for the support recovery problem. This condition is a scaling of the number of measurements $N$, the size of the unknown sparse vector $M$, the sparsity level $k$, and the associated measurement parameters. Under this scaling law the JT decoder recovers the support set asymptotically with probability tends to one. This scaling law can be derived by obtaining the error probability of

support recovery. Deriving an analytical expression for the exact error probability is not a straightforward task. Therefore, instead of analytically obtaining the error probability, an upper bound is derived. In the past works, chi-square tail bounds have been used for bounding the probability of error events [16], [17]. The issue is that the chi-square tail bounds are very loose. A large gap is observed between the exact error probability (obtained for instance by Monte Carlo simulations) and the upper bound derived using the chi-square tail bounds. This large gap will result in a considerable inefficiency in obtaining the required number of measurements for an exact support recovery. It is clear that the number of measurements that is yielded using these bounds is inefficiently more than required. Motivated by this observation, in this paper, we propose a new methodology for obtaining a sharper sufficient condition for an exact support recovery. More details about the proposition are given in Section 3.
We assume the observation model (1), in which $A \in R^{N\times M}$ is the sensing matrix. Actually, the name (sensing matrix) comprises the fact that the signal $x$ is sensed through the matrix $A$ to produce the observation $y$. Elements of $A$ are drawn from standard normal distribution, i.e. $a_{ij} \sim N(0,1)$ in which $a_{ij}$ is the $(i, j)$ element of $A$ and $N(0,1)$ is the normal distribution. The additive noise is $n \sim N(0, \mathrm{I}_{N\times N}\sigma^2)$ where $\mathrm{I}_{N\times N}$ stands for the identity matrix. By $x_i$, we denote the $i$-th entry of vector $x$. The only prior information about the unknown sparse vector $x$ is its sparsity level $k$ that is known at the decoder. We assume that $x$, $A$, and $n$ are statistically independent from each other. A linear regime is considered, in which $M$ and $k$ depend linearly. Also we assume a linear dependency between $k$ and $N$. Two error criteria are considered: zero-one loss for an exact support recovery and a metric for partial support recovery or recovery of most subspace information of $x$. First, an upper bound for the error probability of support recovery is computed, and then the number of sufficient observations for an exact support recovery in the high-dimensional setting $\left((k,N,M)\to\infty\right)$ is derived. It is shown that the sufficient number of observations depends on the noise variance $\sigma^2$, the minimum non-zero entry of the unknown sparse vector $x_{\min}$, and the sparsity level $k$. To support the results further, simulations are provided. Results show that how to choose the upper bound of the error probability is important in obtaining the sufficient number of observations for exact support recovery.
This paper is organized as what follows; Section 2 contains definitions and assumptions. Main results are stated in Section 3. Proofs are given in Section 4. In Section 5, simulations are provided, and conclusions are drawn in Section 6.

## 3. Definitions and assumptions

The following notations are used throughout this paper.
Consider the observation model (1) and the sparsity assumption (2); we refer to the cardinality of $S$ as $|S| = \|x\|_0 = k$, in which $\|x\|_0$ stands for $\ell_0$-norm or the number of non-zero entries of $x$.

*Regime of sparsity*: We assume that there exists a linear dependency between the number of observations and the sparsity level or $M = \alpha k, \alpha \geq 2$.

*Error metrics*: To analyze the performance of the decoder, two error metrics are considered [19], [20]:
$D_{0-1}$: Zero-one error metric:

$$D_{0-1}(x) = 1 - \mathrm{I}\{i \mid \hat{x}_i \neq 0, \forall i \in S\} \times \mathrm{I}\{i \mid \hat{x}_i = 0, \forall i \notin S\} \tag{3}$$

where $\mathrm{I}$ is an indicator function and $\hat{x}$ is the estimation of $x$. This error metric is also known as the exact error metric. This metric declares an error when the estimated support is complement of the true support or when it overlaps with the true support in less than $k$ indices. Let $\hat{S}$ be the estimated support set, which is a *k*-element subset of $\{1,2,...,M\}$. $\hat{S}$ is complement of the true support set $S$ when $S \cap \hat{S} = \Phi$ which means that the estimated support set is completely incorrect. The other error case happens when $S \cap \hat{S} \neq \Phi$ but $S \neq \hat{S}$. It means that $\hat{S}$ overlaps with $S$ in only *q* indices, where *q* < *k*. Again, the estimated support set is not correct. In both of these two cases, the zero-one error metric declares an error.
$D_{MSI}$: This error metric is the statistical extension of the zero-one error metric. It considers recovery of most of the sub-space information of $x$.

$$D_{MSI}(x,\theta)=1-\mathrm{I}\left(\frac{|\{i\,|\,x_i\neq 0\}\cap S|}{|S|}>1-\theta\right) \quad (4)$$

where $\theta\in(0,1)$. This error metric is not as restricted as the zero-one error metric. It allows a small pre-defined amount of distortion. This amount is controlled by $\theta$. When the ratio of incorrectly estimated non-zero positions on $k$ exceeds a pre-defined threshold, this error metric declares an error.

*Sub-matrix and Projection matrix*: A sub-matrix of $A$ containing only columns associated with the index set $S\in\{1,2,...,M\}$ where $|S|=k$ is shown by $A_S$. For any set $S$ with cardinality $k$, we assume $rank(A_S)=k$. The orthogonal projection matrix onto the sub-space spanned by the columns of $A_S$ is shown by $\Pi_{A_S}=A_S\left(A_S^*A_S\right)^{-1}A_S^*$. Also the orthogonal projection matrix onto the orthogonal complement of this sub-space is shown by $\Pi_{A_S}^{\perp}=\mathrm{I}-A_S\left(A_S^*A_S\right)^{-1}A_S^*$.

*Assumption on the decoder*: The decoder is a mapping from pair $(y,A)$ to a set of indices $\hat{S}$. More precisely, it outputs a set of indices $\hat{S}$ with cardinality $k$ as the estimated support. It is assumed that the error probability is averaged over all standard Gaussian measurement matrices with entries confirming $a_{ij}\sim N(0,1)$. The error reads as:

$$P_{eD}=E_A(P(D\neq 0)) \quad (5)$$

in which $P(.)$ stands for the probability measure, $E_A$ is the expectation over all sensing matrices $A$, and $D$ is either $D_{0-1}(x)$ or $D_{MSI}(x,\theta)$.

*JT decoder*: The JT decoder is assumed to be asymptotically optimal. It characterizes events based on their *typicality*. Thus error events are expressed based on *atypicality*. Here, the following definition of the joint-typicality property is exploited:

*Joint-typicality property*: The observation vector $y$ and set $S\in\{1,2,...,M\}$ with $|S|=k$ are $\delta$-jointly typical if $rank(A_S)=k$ and:

$$\left|\frac{1}{N}\left\|\Pi_{A_S}^{\perp}y\right\|^2-\frac{N-k}{N}\sigma^2\right|<\delta \quad (6)$$

The JT decoder outputs an estimate of the support denoted by $\hat{S}$ which is a *k*-element subset of set $\{1,2,...,M\}$. An error is declared when $\hat{S}$:

1. is complement of the true support $S$, and
2. overlaps with $S$ in $0<q<1$ indices.

**3. Main result**

The main idea of the paper can be inferred from Figure 1. In this figure, the exact probability of error and its upper bounds are depicted. The upper bounds 1 and 2 are two different bounds derived using different methodologies. Assume that we want to derive a sufficient number of measurements for which the error probability remains under a specific threshold $\delta$. It is hard to derive the exact value of the error probability for this problem. Thus, in the literature, efforts have been made for deriving acceptable upper bounds rather than computing the exact value of the error probability. In this figure, the exact error probability curve (which may be obtained by simulating the exhaustive search decoder) indicates that for $N_e$ measurements and more the error probability remains under $\delta$. In other words, the $N_e$ measurements are sufficient for a support recovery with an acceptable amount of distortion $\delta$. Since we do not have access to the exact error probability, we have to rely on an upper bound to achieve the sufficient condition. It is clear that the best upper bound among all is the one that is closer to the exact error probability. This illustrative example shows that when a fix distortion level $\delta$ is tolerable, $N_e$ samples are sufficient, whereas using the upper bounds 1 and 2 for deriving a sufficient number of observations imposes the need for additional $(N_1-N_e)$ and $(N_2-N_e)$ samples, respectively. Nonetheless, since $(N_2-N_e)<(N_1-N_e)$, the upper bound 2 is preferable. In other words, to achieve more accuracy and efficiency in the sufficiency proof, we have to exploit an approximation very close to the exact value of the error probability.

In this work, first, an upper bound is derived on the error probability of the support recovery. Compared with the previous bound derived in [20], this upper bound is close to the exact error probability in the high-dimensional setting. In [20], a sufficient condition of support recovery is obtained for the JT decoder in the high-dimensional setting in which the chi-square tail bounds are used to upper bound the probability of error events. The Chi-square tail bounds are very loose for this case [29], [30]. In this paper, we propose a methodology that uses tighter bounds for upper bounding the probability of error events.

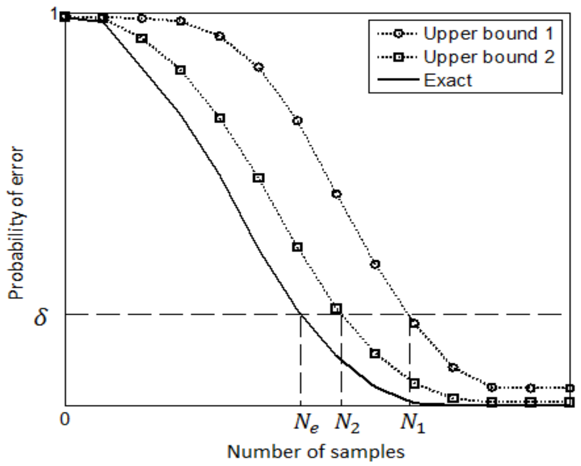


**Figure 1.The error probability versus number of observations. Using a loose upper bound could not be efficient.**

Based on the derived upper bound, a sufficient number of measurements for the exact and approximate support recoveries is derived. It is clear that using the proposed upper bound enhances, the sufficiency proof results. We have assumed that elements of the sensing matrix are *i.i.d*, and they have a normal distribution. The additive noise is also Gaussian. In the literature, some works have been reported considering other types of sensing matrices [13], [18], [29]. In [13], Wang studied the problem considering various types of dense and sparse sensing matrices. Considering the fact that Gaussian measurement matrices are actually highly dense matrices that may lead to prohibitively high computational complexity and storage requirements, he suggested using sparse matrices and tried to find the trade-off between the statistical efficiency and the accuracy. He stated in his paper that the standard Gaussian measurement matrix achieved an optimal scaling of the number of observations required for the support recovery.

In what follows, the main results are presented in theorems 1 and 2. In theorem 1, a new upper bound is derived on the error probability of the support recovery. Using theorem 1, in theorem 2, a sufficient condition for exact support recovery in the high-dimensional setting is derived.

### 3.1. Error events

A specified form of the JT decoder is used and analyzed for signal recovery. The JT decoder characterizes the events based on their typicality. Thus error events are expressed based on atypicality. Consider the following two events:

$$E_S = \{y \text{ and } S \text{ are } \delta\text{-jointly typical}\}$$
$$E_S^C = \{\text{occurance of complement of } S\}$$

Let $\hat{S} \neq S$ be the estimated support such that $|\hat{S}| = k$, $|\hat{S} \cap S| = q < k$, and $rank(A_{\hat{S}}) = k$. Event $E_S^C$ implies occurrence of the complement of the true support and event $E_{\hat{S}}$ occurs when $y$ and $\hat{S}$ are $\delta-$jointly typical or when $S$ and $\hat{S}$ overlaps in $q < k$ indices. For each $q$, there are $N(q)$ number of subsets $\hat{S}$ with the mentioned properties, where:

$$N(q) = \binom{k}{q}\binom{M-k}{q} \tag{7}$$

If $E_S^C$ or $E_{\hat{S}}$ occurs, then the decoder fails. The probability of error can be expressed as:

$$P_{eD} = P\left(E_S^C \cup \left(\bigcup_{\{|\hat{S}|=k, \hat{S}\neq S\}} E_{\hat{S}}\right)\right) \leq P\left(E_S^C\right) + \sum_{q=1}^{k} N(q) P\left(E_{\hat{S}}\right) \tag{8}$$

To derive the upper bound on the error probability $P_{eD}$, it suffices to compute probability of events $E_S^C$ and $E_{\hat{S}}$. By multiplying both sides of the observation model $y = Ax + n$ by $\Pi_{A_S}^{\perp}$, we get $\Pi_{A_S}^{\perp} y = \Pi_{A_S}^{\perp} n$. Using the substitution $\Pi_{A_S}^{\perp} = U_S \Lambda U_S^*$, in which asterisk stands for the conjugate transposed of the corresponding matrix, $U_S$ is a unitary matrix and $\Lambda$ is a diagonal matrix with the first $N-k$ diagonal entries equal to 1 and the remaining entries equal to zero, we obtain:

$$\begin{aligned}\left\|\Pi_{A_S}^{\perp} y\right\|^2 &= \left\|U_S \Lambda U_S^* \left(Ax+n\right)\right\|^2 \\ &= \left\|U_S \Lambda U_S^* n\right\|^2 = \left\|\Lambda U_S^* n\right\|^2 = \left\|\Lambda n'\right\|^2 \\ &= \left|n_1'\right|^2 + \left|n_2'\right|^2 + \ldots + \left|n_{N-k}'\right|^2\end{aligned}$$

in which each element of vector $n' = U_S^* n$ is an *i.i.d* random variable with distribution $N(0,\sigma^2)$. Thus $G = \frac{1}{\sigma^2}\left\|\Pi_{A_S}^{\perp} y\right\|^2$ is a chi-square random variable with $N-k$ degrees of freedom.

Again, by multiplying both sides of the observation model by $\Pi_{A_{\hat{S}}}^{\perp}$, we get $\Pi_{A_{\hat{S}}}^{\perp} y = \Pi_{A_{\hat{S}}}^{\perp}\left(Ax+n\right)$. By substitution $\Pi_{A_{\hat{S}}}^{\perp} = U_{\hat{S}} \Lambda U_{\hat{S}}^*$, in which $U_{\hat{S}}$ is a unitary matrix and the first $N-k$ diagonal entries of diagonal matrix $\Lambda$ are equal to 1 and the remaining entries are equal to zero, we get:

$$\begin{aligned}\left\|\Pi_{A_{\hat{S}}}^{\perp} y\right\|^2 &= \left\|U_{\hat{S}} \Lambda U_{\hat{S}}^* \left(\sum\nolimits_{i\in S\backslash \hat{S}}\left(x_i a_i + n\right)\right)\right\|^2 \\ &= \left\|\Lambda\left(\sum\nolimits_{i\in S\backslash \hat{S}}\left(x_i U_{\hat{S}}^* a_i + U_{\hat{S}}^* n\right)\right)\right\|^2 \\ &= \left|z_1\right|^2 + \left|z_2\right|^2 + \ldots \left|z_{N-k}\right|^2\end{aligned}$$

where, $a_i$ is the $i$-th column of matrix $A$. For all $i \in S \backslash \hat{S}$, vector $a_i' = U_{\hat{S}}^* a_i$ has *i.i.d* entries each with distribution $N(0,1)$, and vector $n' = U_{\hat{S}}^* n$ has *i.i.d* entries, each with distribution $N(0,\sigma^2)$. Consequently, $z_i$ has also *i.i.d* entries each with distribution $N(0,\sigma_1^2)$ such that $\sigma_1^2 = \sum_{q\in S\backslash \hat{S}}\left|x_q\right|^2 + \sigma^2$ Again, $\hat{G} = \frac{1}{\sigma_1^2}\left\|\Pi_{A_{\hat{S}}}^{\perp} y\right\|^2$ is a chi-square random variable with $N-k$ degrees of freedom. For $\delta > 0$, consider the following two error probabilities:

$$P\left(E_S^C\right) = P\left(\left|\frac{1}{N}\left\|\Pi_{A_S}^{\perp} y\right\|^2 - \frac{N-k}{N}\sigma^2\right| > \delta\right) \tag{9}$$

$$P\left(E_{\hat{S}}\right) = P\left(\left|\frac{1}{N}\left\|\Pi_{A_{\hat{S}}}^{\perp} y\right\|^2 - \frac{N-k}{N}\sigma^2\right| < \delta\right) \tag{10}$$

We have:

$$\begin{aligned}P_{eD} &= P\left(\hat{S} \neq S\right) \leq P\left(E_S^C\right) \\ &= P\left(\left|\frac{1}{N}\left\|\Pi_{A_S}^{\perp} y\right\|^2 - \frac{N-k}{N}\sigma^2\right| > \delta\right) \\ &+ \sum_{q=h}^{k} N(q) P\left(\left|\frac{1}{N}\left\|\Pi_{A_{\hat{S}}}^{\perp} y\right\|^2 - \frac{N-k}{N}\sigma^2\right| < \delta\right)\end{aligned} \tag{11}$$

in which $h \in \left[1, \theta k\right]$. When $h=1$, the zero-one error metric is considered. As it can be inferred from (11), $h=1$ means that the error is computed for all non-zero positions that are estimated incorrectly, and in the summation, their appropriate error events are taken into account, whereas $h=\theta k$ means that the error is computed when the number of non-zero positions that are incorrectly estimated exceeds a certain amount, $\theta k$. The first term in the right hand side is probability of event $E_S^C$, which is occurrence of the complement of the true support set. The second term comprises all $E_{\hat{S}}$ events that occur when $y$ and $\hat{S}$ are $\delta-$jointly typical or when $S$ and $\hat{S}$ overlap in $q < k$ indices. $\delta$ is the distortion parameter. Here, we have assumed that $\delta < \min\left\{x_{\min}^2, \sigma^2\right\}$.

### 3.2. Results

*Theorem 1*- From the $k$-sparse signal $x \in R^M$ with the support (2), a linear noisy observation $y = Ax+n$ is generated. Elements of $A \in R^{N\times M}$ are drawn from standard Gaussian distribution $a_{ij} \sim N(0,1)$, and $n \sim N(0, \mathrm{I}_{N\times N}\sigma^2)$ is the additive Gaussian noise. Assume that $rank(A_S) = k$ and for any set $T \subset \left\{1,2,\ldots,M\right\}$ with $\left|T\right| = k$, let $P(rank(A_T) < k) = 0$. For $\delta > 0$ the probability of support recovery error is bounded above according to the following equation:

$$P_{eD} < U \tag{12}$$

where, $U = A_1 + A_2 + A_3$, and $A_1$, $A_2$ and $A_3$ are as follow:

$$A_1 = \frac{\exp\left(\frac{\delta N}{2\sigma^2}\right)\left(\frac{N-k}{2}.\frac{\delta N}{2\sigma^2}\right)^{\frac{N-k}{2}}}{\mathrm{e}\sqrt{\pi}G\left(\frac{N-k}{2}-1\right)^{\frac{N-k}{2}-1}\left(\frac{\delta N}{2\sigma^2}\right)} \quad (13)$$

$$A_2 = \frac{B\left(\frac{N-k}{2}+\frac{\delta N}{2\sigma^2}\right)^{\frac{N-k}{2}-1}\exp\left(-\frac{\delta N}{2\sigma^2}\right)}{e\sqrt{\pi}G\left(\frac{N-k}{2}-1\right)^{\frac{N-k}{2}-1}} \quad (14)$$

$$A_3 = \sum_{q=h}^{k} N(q)\frac{\exp\left(-\frac{(N-k)\sigma^2+\delta N}{2\sigma_{1,\min}^2}+\frac{N-k}{2}-1\right)\left(\frac{(N-k)\sigma^2+\delta N}{2\sigma_{1,\min}^2}\right)^{\frac{N-k}{2}}}{\left(\frac{N-k}{2}-\frac{(N-k)\sigma^2+\delta N}{2\sigma_{1,\min}^2}\right)\sqrt{\pi}G\left(\frac{N-k}{2}-1\right)^{\frac{N-k}{2}-1}} \quad (15)$$

in which:

$$G=\left(8\left(\frac{N-k}{2}-1\right)^3+4\left(\frac{N-k}{2}-1\right)^2+\left(\frac{N-k}{2}-1\right)+\frac{1}{100}\right)^{\frac{1}{6}},$$

$$\sigma_{1,\min}^2 = qx_{\min}^2+\sigma^2,$$

and $x_{\min} = \min_{i\in S}|x_i|$. $h$ is 1 for zero-one error metric and $\theta k$ for error metric 2.

*Theorem 2*- From the *k*-sparse signal $x \in R^M$ with support (2), a linear noisy observation $y = Ax + n$ is generated. Elements of $A \in R^{N\times M}$ are drawn from standard Gaussian distribution $a_{ij} \sim N(0,1)$, and $n \sim N(0, \mathrm{I}_{N\times N}\sigma^2)$ is additive Gaussian noise. Assume that $rank(A_S) = k$ and for any $T \subset \{1,2,...,M\}$ with cardinality *k*, let $P(rank(A_T) < k) = 0$. The sufficient condition for an exact support recovery in the high-dimensional setting is as follows:

$$N > k\max\left\{\left(\frac{\sigma^2}{\sigma^2-\delta}\right),\left(\frac{x_{\min}^2}{x_{\min}^2-\delta}\right)\right\}. \quad (16)$$

In fact, theorem 2 gives the condition on the number of observations such that $P_{eD} \to 0$ as $k$ and consequently, $M$ and $N$ grow large.

## 4. Proofs

In what follows, two sub-sections, proofs and further discussions, are provided.

### 4.1.Proof of theorem 1

As it was mentioned in Section 3, $G = \frac{1}{\sigma^2}\left\|\Pi_{A_S}^{\perp}y\right\|^2$ is a chi-square random variable with $N-k$ degrees of freedom. Using the probability distribution function of a chi-square random variable, we have:

$$\begin{aligned} P\left(E_S^C\right) &= P\left(\left|\frac{1}{N}\left\|\Pi_{A_S}^{\perp}y\right\|^2-\frac{N-k}{N}\sigma^2\right|>\delta\right) \\ &= P\left(\left|G-(N-k)\right|>\frac{\delta N}{\sigma^2}\right) \\ &= P\left(G-(N-k)>\frac{\delta N}{\sigma^2}\right) \\ &\quad + P\left(G-(N-k)<-\frac{\delta N}{\sigma^2}\right) \\ &= \frac{1}{\Gamma(z)}\int_0^{\frac{c_1}{2}} t^{z-1}\exp(-t)\,dt \\ &\quad + \frac{1}{\Gamma(z)}\int_{\frac{c_2}{2}}^{\infty} t^{z-1}\exp(-t)\,dt \end{aligned} \quad (17)$$

in which:

$$z=\frac{N-k}{2},$$

$$c_1 = N-k-\frac{\delta N}{\sigma^2},$$

$$c_2 = N-k+\frac{\delta N}{\sigma^2}.$$

Since we are assuming that $c_1 > 0$, we have:

$$\frac{N}{k} = \beta > \frac{\sigma^2}{\sigma^2-\delta} \quad (18)$$

We put an upper bound on the right hand side of (17) and show that it tends to zero as the problem dimensions grow large. We know that the first and second integrals in (17) are the lower and upper incomplete gamma functions, respectively.

We use the following lemmas, which applies to any gamma, lower incomplete gamma, and upper incomplete gamma functions to upper bound the integrals [25-27].

*Lemma 1*. Let $u \geq 1$ and real; then:

$$\sqrt{\pi}\frac{u^u}{\exp(u)}\left(8u^3+4u^2+u+\frac{1}{100}\right)^{\frac{1}{6}} < \Gamma(u+1) < \sqrt{\pi}\frac{u^u}{\exp(u)}\left(8u^3+4u^2+u+\frac{1}{30}\right)^{\frac{1}{6}} \tag{19}$$

*Lemma 2*. For $\frac{c_1}{2} < z$, the following inequality holds:

$$\int_0^{\frac{c_1}{2}} t^{z-1}\exp(-t)\,dt < \frac{\exp\left(-\frac{c_1}{2}\right)\left(\frac{c_1}{2}\right)^z}{z-\frac{c_1}{2}} \tag{20}$$

*Lemma 3*. For $x>1$, $B>1$ and $y>\frac{B}{B-1}(x-1)$ the following inequality holds:

$$y^{x-1}\exp(-y) < \left|\Gamma(x,y)\right| < By^{x-1}\exp(-y) \tag{18}$$

The first integral in the right hand side of (17) can be bounded using the lower bound of the gamma function in (19) and the upper bound of the lower incomplete gamma function of (20). The second integral can be bounded using the lower bound of the gamma function in (19), and the upper bound of the upper incomplete gamma function in (21). Thus we get:

$$\begin{aligned} &P\left(E_S^C\right) < \\ &\frac{\left(\frac{\exp\left(-\frac{c_1}{2}+z-1\right)\left(\frac{c_1}{2}\right)^z}{z-\frac{c_1}{2}}\right)+B\left(\frac{c_2}{2}\right)^{z-1}\exp\left(-\frac{c_2}{2}+z-1\right)}{\sqrt{\pi}G(z-1)^{z-1}} \\ &= A_1 + A_2 \end{aligned} \tag{22}$$

Similarly, we know that $\hat{G}=\frac{1}{\sigma_1^2}\left\|\Pi_{A_{\hat{S}}}^{\perp}y\right\|^2$ is a chi-square random variable with $N-k$ degrees of freedom. Thus we have:

$$\begin{aligned} P\left(E_{\hat{S}}\right) &= P\left(\left|\frac{1}{N}\left\|\Pi_{A_{\hat{S}}}^{\perp}y\right\|^2-\frac{N-k}{N}\sigma^2\right|<\delta\right) \\ &= P\left(\left|\hat{G}-(N-k)\frac{\sigma^2}{\sigma_1^2}\right|<\frac{\delta N}{\sigma_1^2}\right) \\ &= P\left(\hat{G}-(N-k)\frac{\sigma^2}{\sigma_1^2}<\frac{\delta N}{\sigma_1^2}\right) \\ &= \frac{1}{\Gamma(z)}\int_0^{\frac{c_3}{2}} t^{z-1}\exp(-t)\,dt \end{aligned} \tag{23}$$

where $c_3 = \frac{(N-k)\sigma^2}{\sigma_1^2}+\frac{\delta N}{\sigma_1^2}$. Using lemma 1 and lemma 2, we get:

$$\begin{aligned} P\left(E_{\hat{S}}\right) &< \frac{\exp\left(-\frac{c_3}{2}+z-1\right)\left(\frac{c_3}{2}\right)^z}{\sqrt{\pi}G(z-1)^{z-1}\left(z-\frac{c_3}{2}\right)} \\ &< \frac{\exp\left(-\frac{(N-k)\sigma^2+\delta N}{2\sigma_{1,\min}^2}+z-1\right)\left(\frac{(N-k)\sigma^2+\delta N}{2\sigma_{1,\min}^2}\right)^z}{\sqrt{\pi}G(z-1)^{z-1}\left(z-\frac{(N-k)\sigma^2+\delta N}{2\sigma_{1,\min}^2}\right)} \\ &= A_3 \end{aligned}$$

and consequently:

$$\begin{aligned} P_{eD} &< \frac{\left(\frac{\exp\left(\frac{c_1}{2}+\frac{N-k}{2}-1\right)\left(\frac{c_1}{2}\right)^z}{z-\frac{c_1}{2}}\right)+B\left(\frac{c_2}{2}\right)^{z-1}\exp\left(-\frac{c_1}{2}+\frac{N-k}{2}-1\right)}{\sqrt{\pi}G\left(\frac{N-k}{2}-1\right)^{\frac{N-k}{2}-1}} \\ &+\sum_{q=h}^{k}N(q)\frac{\exp\left(-\frac{(N-k)\sigma^2+\delta N}{2\sigma_{1,\min}^2}+\frac{N-k}{2}-1\right)\left(\frac{(N-k)\sigma^2+\delta N}{2\sigma_{1,\min}^2}\right)^z}{\sqrt{\pi}G\left(\frac{N-k}{2}-1\right)^{\frac{N-k}{2}-1}\left(z-\frac{(N-k)\sigma^2+\delta N}{2\sigma_{1,\min}^2}\right)} \end{aligned} \tag{24}$$

where, for the zero-one error metric and error metric 2, $h$ is 1 and $\theta k$, respectively.

### 4.2. Proof of theorem 2

In order to prove that $P_{eD} \to 0$ asymptotically, it is sufficient to show that the upper bound of (24) tends to zero as $k$ grows large. Since the linear regime is assumed, we have $M=\alpha k, (\alpha>2)$ and also we assume that $N=\beta k,(\beta>1)$. It is straightforward to show that $A_1$ and $A_2$ tend to zero asymptotically. We write:

$$\frac{\left(\frac{\exp\left(-\frac{c_1}{2}+z-1\right)\left(\frac{c_1}{2}\right)^z}{z-\frac{c_1}{2}}\right)+B\left(\frac{c_2}{2}\right)^{z-1}\exp\left(-\frac{c_2}{2}+z-1\right)}{\sqrt{\pi}G(z-1)^{z-1}}=$$
$$\frac{\exp\left(\frac{\delta\beta}{2\sigma^2}k\right)\left(\left(\frac{\beta-1}{2}\right)k-1\right)}{\sqrt{\pi}\frac{\delta\beta}{2\sigma^2}k\left(\frac{\beta-1}{\beta-1-\frac{\delta\beta}{\sigma^2}}\right)^{\left(\frac{\beta-1}{2}\right)k}\left(8\left(\left(\frac{\beta-1}{2}\right)k-1\right)^3+4\left(\left(\frac{\beta-1}{2}\right)k-1\right)^2+\left(\left(\frac{\beta-1}{2}\right)k-1\right)+\frac{1}{100}\right)^{\frac{1}{6}}}+$$
$$\frac{B\exp\left(-\frac{\delta\beta}{2\sigma^2}k\right)\left(1+\frac{\delta\beta}{\sigma^2(\beta-1)}\right)^{\left(\frac{\beta-1}{2}\right)k}\left(\frac{\frac{(\beta-1)}{2}k-1}{\left(\frac{(\beta-1)}{2}-\frac{\delta}{2\sigma^2}\beta\right)k}\right)}{\sqrt{\pi}\left(8\left(\left(\frac{\beta-1}{2}\right)k-1\right)^3+4\left(\left(\frac{\beta-1}{2}\right)k-1\right)^2+\left(\left(\frac{\beta-1}{2}\right)k-1\right)+\frac{1}{100}\right)^{\frac{1}{6}}} \tag{25}$$

where, $M$ and $N$ are written in terms of $k$. Using (18), one can show that the first term in the right hand side converges. For the second term to converge, it suffices to show:

$$\left(1+\frac{\delta\beta}{\sigma^2(\beta-1)}\right)^{\left(\frac{\beta-1}{2}\right)}\exp\left(-\frac{\delta\beta}{2\sigma^2}\right)<1 \tag{26}$$

Since we assume $\beta>\frac{\sigma^2}{\sigma^2-\delta}$, inequality (26) is true using the fact that $\log(1+x)-x<0$ for all $x>-1$. Next, it remains to show that $A_3$ tends to zero asymptotically. Let $h=1$. Using the following bound:

$$\binom{k}{q}<\exp\left(q\log\left(\frac{ke}{q}\right)\right) \tag{27}$$

we have:

$$A_3<\sum_{q=1}^{k}\exp\left(q\log\left(\frac{ke}{q}\right)+q\log\left(\frac{(M-k)e}{q}\right)\right)$$
$$\exp\left(\frac{\exp\left(-\frac{(N-k)\sigma^2+\delta N}{2\sigma_{1,\min}^2}\right)\left(\frac{(N-k)\sigma^2+\delta N}{2\sigma_{1,\min}^2}\right)^z}{\sqrt{\pi}G\left(e^{-1}\frac{N-k}{2}-e^{-1}\right)^{\frac{N-k}{2}-1}\left(z-\frac{(N-k)\sigma^2+\delta N}{2\sigma_{1,\min}^2}\right)}\right) \tag{28}$$

To get rid of summation, we obtain the maximizer of (28) in terms of $q$. We replace $q$ with a continuous counterpart $x\in[1,k]$. Consider the function below, which is part of $A_3$ that contains $x$:

$$g(x)=\frac{\exp\left(2x\left(\log\left(\frac{a}{x}\right)-\frac{b}{x+c}\right)\right)\left(\frac{b}{x+c}\right)^z}{\left(z-\frac{b}{x+c}\right)} \tag{29}$$

in which:

$$a=ke\sqrt{\beta-1},b=\frac{(N-k)\sigma^2+\delta N}{2x_{\min}^2},c=\frac{\sigma^2}{x_{\min}^2}$$

One can show that $g(x)$ is strictly ascending when:

$$\beta>\frac{x_{\min}^2}{x_{\min}^2-\delta} \tag{30}$$

$A_3$ reaches its maximum value when $q=k$. Therefore, it is easy to show that the right hand side of (28) tends to zero as $k$ tends to infinity. For error metric 2, the expression must tend to zero for $k$ larger than $\theta h$. This result is also applied for the error metric 2. Considering (18) and (30), the sufficient condition is derived.

In the previous similar work [20], an assumption was considered for $x_{\min}^2$, where:

$$\frac{kx_{\min}^2}{\log(k)}\to\infty \quad as \quad k\to\infty$$

The reason for accepting this condition is that in the noisy setting, when elements of $x$ are arbitrarily small, a perfect recovery is not possible. Furthermore, condition (30) does not conflict with this outcome since one can result this condition from (30).

## 5. Simulation results

In this section, a simulation is provided, which shows the efficiency of the proposed method. Settings of the problem parameters is such that condition (18) is satisfied. A comparison is made between our results and the previous similar work [20]. The upper bound of the error probability derived in [20] is denoted by $V$. In Figure 2, $U$ is the proposed upper bound. For the zero-one error metric $U$ is depicted as a function of $k$ for three different values of $\beta$ and it is compared with $V$. It is shown that the convergence rate of $U$ is more than that of $V$. As it is expected, for $k$ sufficiently large, $U$ is a sharper upper bound for the error probability compared with $V$. To achieve a pre-defined distortion level, a stricter upper bound results in obtaining fewer number of measurements for the support recovery. Obviously, using a sharp upper bound does lead to a considerable improvement in the sufficient condition. Another worth mentioning issue that can be inferred from Figure 2 is that $U$ shows more sensitivity to changing $\beta$ than $V$. A very small increase in $\beta$ decreases $U$ somehow but has no significant effect on $V$. Sensitivity to $\beta$ which means sensitivity to the number of measurements, is an expected property for a good upper bound. Thus $U$ provides a more exact value for $\beta$ than $V$ does. In Figure 3, $U$ and $V$ are plotted versus $k$ while $\sigma^2$ decreases. Decreasing $\sigma^2$ makes both $U$ and $V$ to decrease, which is expected. Again, it can be seen that sensitivity of $U$ to changing $\sigma^2$ is more than that of $V$. This featuree is not an improvement itself but it is not very important since for a sufficiently large $k$, $U$ is negligible against $V$.

Totally, the simulation results show that using the proposed approach enhances the sufficient condition for the support recovery. In fact, the previous upper bound $V$ imposes the need for some additional measurements. Since this bound is derived based on chi-square tail bounds and it is very loose, it cannot provide a near to exact sufficient condition (required number of measurements). It can only give an approximation of the sufficient number of measurements for sparsity recovery. However, as it is confirmed by simulations, since the proposed upper bound $U$ is very close to the exact error probability, it results in increasing the accuracy in sufficiency proof.

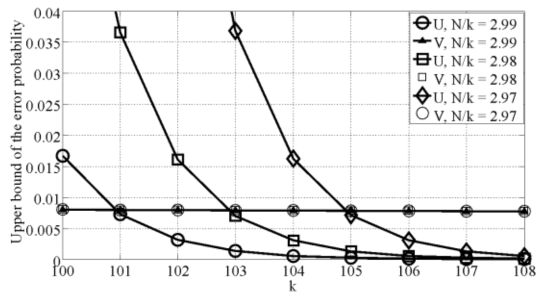


**Figure 2. Upper bound of the error probability as a function of $k$ for three different values of $\beta$. It is shown that the convergence rate of U and also its sensitivity to changing $\beta$ are more than those of V. For a desired distortion level, using U as the upper bound of the error probability results in a smaller $\beta$ than using V.**

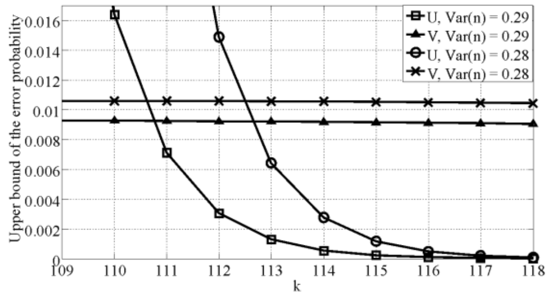


**Figure 3. Upper bound of the error probability as a function of $k$ for two different values of $\sigma^2$.**

## 6. Conclusion

In this work, we examined the sufficient condition for the sparsity pattern recovery. The analyses were based on a joint-typicality decoder. Considering the linear regime, when the sensing matrix contained *i.i.d.* normal random entries and the noise was Gaussian, we computed an upper bound on the probability of error. It was shown analytically and also using simulations that the derived upper bound was tighter than the previous loose upper bounds, which were derived based on the chi-square tail bounds. Based on the proposed upper bound, a sufficient number of measurements for an exact sparsity pattern recovery was obtained. It was shown that the sufficient number of measurements for an exact support recovery depends on the noise variance, the minimum nonzero entry of the unknown sparse vector, and the sparsity level, and it was shown to improve the existing results [20]. This work can further be extended to arbitrary sensing matrices.

## References

[1] Donoho, D. L. (2006). Compressed sensing, IEEE. Trans. Inf. Theory., vol. 52, no. 4, pp. 1289-1306.

[2] Chen, S., et al. (2003). Atomic decomposition by basis pursuit, SIAM J. Sci. Comput., vol. 20, pp. 33-61.

[3] Candes, E. J., et al. (2006). Stable signal recovery for incomplete and inaccurate measurements, Commun. Pure Appl. Math., vol. 59, pp. 1207-1223.

[4] Elad, M. (2010). Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing, Springer.

[5] Abolghasemi, V. (2015). A block-wise random sampling approach: Compressed sensing problem, Journal of AI and Data Mining, Shahrood University, vol. 3, pp. 93-100.

[6] Du, X., et al. (2013). A reduced $l_2 - l_1$ model with an alternating minimization algorithm for support recovery of multiple measurement vectors, IET Signal Process., vol. **7**, no. 2, pp. 112-119.

[7] Baillet, S., et al. (2001). Electromagnetic brain mapping, IEEE Signal Process. Mag., vol. 18, no. 6, pp. 14-30.

[8] Wipf, D. & Nagarajan, S. (2008). A unified Bayesian framework for meg/eeg source imaging, NeuroImage, vol. 44, no. 3, pp. 947-966.

[9] Tian, Z. & Giannakis, G. B. (2007). Compressed sensing for wideband cognitive radios. Proc. Int. Conf. Acoustics, Speech and Signal Processing, Honolulu, Hl, pp. 1357–1360.

[10] A. J. Miller, “Subset selection in regression”, Chapman&Hall, 2002.

[11] Guo, D. (2009). Neighbour discovery in ad hoc networks as a compressed sensing problem, Presented at Information Theory and Application Workshop, UCSD.

[12] Akcakaya, M. & Tarokh, V. (2008). Noisy compressive sampling limits in linear and sub-linear regimes, Proc. Conf. Information Sciences and Systems, pp. 1-4.

[13] Wang, W. & Wainwright, M. J. (2010). Information-theoretic limits on sparse signal recovery: Dense versus sparse measurement matrices, IEEE Trans. Inf. Theory., vol. 56, no. 6, pp. 2967-2977.

[14] Aeron, S., et al. (2010). Information theoretic bounds for compressed sensing, IEEE Trans. Inf. Theory., vol. 56, no. 10, pp. 5111-5130.

[15] Wainwright, M. J. (2009). Sharp thresholds for high-dimensional and noisy sparsity recovery using $l_1$-constrained quadratic programming (Lasso), IEEE Trans. Inf. Theory., vol. 55, no. 5, pp. 2183-2202.

[16] Scarllet, J., et al. (2013). Compressed sensing with prior information: information-theoretic limits and practical decoders, IEEE. Trans. Information theory, vol. 61, no. 2, pp. 427-439.

[17] Park, S. & Lee, H. N. (2012). Number of compressed measurements needed for noisy distributed compressed sensing, IEEE Inter. Symp. Information Theory Proceedings. pp. 1648-1651.

[18] Jin, Y., et al. (2011). Limits on support recovery of sparse signals via multiple-access communication techniques, IEEE Trans. Inf. Theory., vol. 57, np. 12, pp. 7877–7892.

[19] Wainwright, M. J. (2009). Information-theoretic limits on sparsity recovery in the high dimensional and noisy setting, IEEE. Trans. Information theory, vol. 55, no. 22, pp. 5728-5741.

[20] Ackakaya, M. & Tarokh, V. (2010). Shannon-theoretic limit on noisy compressive sampling, IEEE. Trans. Information theory, vol. 56, no. 1, pp. 492-504.

[21] Reeves, G. & Gastpar, M. (2013). Approximate sparsity pattern recovery: Information theoretic lower bounds, IEEE. Trans. Information theory., vol. 59, no. 6, pp. 3451-3465.

[22] Fletcher, A. K., et al. (2009). Necessary and sufficient conditions for sparsity pattern recovery, IEEE. Trans. Information theory, vol. 55, no. 12, pp. 5758-5772.

[23] Xu, W., et al. (2014). Performance analysis of partial support recovery and signal reconstruction of compressed sensing, IET Signal. Process, vol. 8, no. 2, pp. 188-201.

[24] Rad, K. (2011). Nearly sharp sufficient conditions on exact sparsity pattern recovery, IEEE Trans. Inf.Theory., vol. 57, no. 7, pp. 4672-4679.

[25] Alzer, H. (2003). On Ramanujan’s double inequality for the gamma function, vol. 35, pp. 601-607.

[26] Merkle, M. (1993). Some inequality for chi-square distribution function and the exponential function, Arc. Math., vol. 60, pp. 451-458.

[27] Natalini, P. & Palumbo, B. (2000). Inequalities for the incomplete gamma function, Math. Inequal. Appl., vol. 3, pp. 69-77.

[28] Lauren, B. & Massart, P. (1998). Adaptive estimation of a quadratic functional by model selection, Ann. Statist., vol. 28, no. 5, pp. 1303-1338.

[29] Shaeiri, Z., et al. (2017). Enhancing the fundamental limits of sparsity pattern recovery, DSP Elsevier, vol. 69, pp. 275-285.

[30] Shaeiri, Z., et al. (2018). Enhancing the sufficient condition of sparsity pattern recovery, DSP Elsevier, vol. 81, pp. 43-49.

[31] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso, J. Roy. Statist. Soc., Series B (Methodol.), vol. 58, no. 1, pp. 267–288.

[32] Chen, S. (1998). Atomic decomposition by basis pursuit, SIAM J. Sci. Comput., vol. 20, pp. 33–61.

[33] Dai, W. & Milenkovic, O. (2009). Subspace pursuit for compressive sensing signal reconstruction, IEEE Trans. Inf. Theory, vol. 55, no. 5, pp. 2230–2249.

[34] Needell, D. & Tropp, J. A. (2009). CoSaMP: Iterative signal recovery from incomplete and inaccurate samples, App. Comput. Harmon. Anal., vol. 26, no. 3, pp. 301–321.

[35] Blumensath, T. & Davies, M. E. (2009). Iterative hard thresholding for compressed sensing, App. Comput. Harmon. Anal., vol. 27, no. 3, pp. 265–274.

نشریه هوش مصنوعی و داده کاوی
IAIDM

# محاسبه شرط کافی دقیق‌تر در بازیابی تنکی

**زهرا شعیری*، محمدرضا کرمی و علی آقاگل‌زاده**

**دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی نوشیروانی بابل، بابل.**

ارسال ۲۰۱۸/۰۳/۱۸؛ بازنگری ۲۰۱۸/۰۴/۲۶؛ پذیرش ۲۰۱۹/۰۶/۰۲

**چکیده:**

در این مقاله یک شـرط کافی برای بازیابی تنکی سـیگنال با در اختیار داشـتن مشـاهده‌ای خطی و نویزی در ابعاد بالا ارائه می‌شـود. گرچه این مسأله در تحقیقات گذشــته مورد توجه قرار گرفته اسـت ولی نتایج می‌توانند بهبود یابند. در این مقاله روشــی برای کاهش دادن شکاف میان کران بالای موجود بر احتمال خطای بازیابی تنکی و مقدار دقیق این احتمال پیشـنهاد می‌شود. این امر به محاسبه شرط کافی دقیق‌تری در بازیابی تنکی منجر می‌شود. جهت بازیابی تنکی از نوع خاصـی از رمزگشای jointly-typical استفاده می‌شود و از دو معیار صحت سنجی بازیابی دقیق و بازیابی پاره‌ای جهت محاسبه خطا اسـتفاده می‌شـود. ابتدا یک کران بالا برای احتمال خطای بازیابی تنکی محاسـبه می‌شـود. سپس براساس کران بالای محاسبه شده یک شرط کافی برای بازیابی تنکی اسـتخراج می‌شود. نشان داده می‌شود که شرط کافی برای بازیابی تنکی به سه پارامتر کلیدی مسأله بستگی دارد؛ واریانس نویز، کوچکترین درایه نامنفی بردار تنک مجهول و سـطح تنکی. نتایج به دسـت آمده به طور تئوری اثبات شـده و در کنار آن شـبیه‌سازی نیز جهت پشتیبانی نتایج ارائه شـده اسـت. شبیه‌سازی‌ها برای نرخ تنکی‌های مختلف و واریانس نویزهای مختلف انجام شده است. نتایج نشان دهنده دقیق‌تر بودن شرط کافی محاسبه شده در مقایسه با شرط کافی پیشین می‌باشند.

**کلمات کلیدی:** بازیابی تنکی، رمزگشای jointly-typical، حسگری فشرده، شرط کافی بازیابی تنکی، محدوده‌های تئوری اطلاعاتی.