# Improved Coherence Index-Based Bound in Compressive Sensing

Ljubiša Stanković, *Fellow, IEEE*, Miloš Brajović, *Member, IEEE*, Danilo Mandic, *Fellow, IEEE*, Isidora Stanković, *Member, IEEE*, and Miloš Daković, *Member, IEEE*

***Abstract*—Within the compressive sensing (CS) paradigm, sparse signals can be reconstructed based on a reduced set of measurements, whereby reliability of the solution is determined by its uniqueness. With its mathematically tractable and feasible calculation, the coherence index is one of very few CS uniqueness metrics with considerable practical importance. We propose an improvement of the coherence-based uniqueness relation for the matching pursuit algorithms. Starting from a simple and intuitive derivation of the standard uniqueness condition, based on the coherence index, we derive a less conservative coherence index-based lower bound for signal sparsity. The results are generalized to the uniqueness condition of the $l_0$-norm minimization for a signal represented in two orthonormal bases.**

***Index Terms*—Compressive sensing, Signal reconstruction, Data acquisition, OMP.**

## I. Introduction and Basic CS Setting

COMPRESSIVE Sensing (CS) is a field that provides a framework for efficient data acquisition [1]–[5], [11]. Examples include applications that rest upon reliable sensing from the lowest possible number of measurements, such as the recovery of sparse signals from vastly reduced sets of measurements and practical solutions in critical cases when some measurements are physically unavailable or heavily corrupted by disturbance.

Within the CS theory, several approaches have been established to reconstruct a sparse, $N$-dimensional vector, $\mathbf{X}$, from a reduced $M$-dimensional set of measurements, $\mathbf{y}$. The main concern in the reconstruction is to provide the conditions for a unique solution. Several frameworks for establishing the conditions for a unique solution have been developed. The most important ones rely on the restricted isometry property (RIP) and the coherence index. While the RIP-based approach provides theoretically well-founded conditions, its computational complexity remains the main problem [1], [2], [12], [13]. Namely, the RIP constant calculation is even more computationally demanding than the direct combinatorial solution of the CS problem itself. On the other hand, the coherence index-based condition is simple and computationally efficient, while its main disadvantage is that the reconstruction conditions based on this metric are quite pessimistic [4], [6].

Here, we introduce an approach which alleviates this deficiency of the coherence index; this is achieved through a computationally simple improved bound for the uniqueness relation based on the coherence index. The approach will be applied to a signal representation in two bases [14], being used for the derivation of the general sparsity bounds when the $\ell_0$ and $\ell_1$ minimizations are used to solve a CS problem.

### A. Definitions and Notation

A sequence $\{X(k)\}$, $k = 0, 1, \ldots, N-1$ is referred to as a sparse sequence if the number, $K$, of its nonzero elements, $X(k) \neq 0$, is much smaller than its total length, $N$, that is,

$$X(k) \neq 0 \text{ for } k \in \{k_1, k_2, \ldots, k_K\}, \ K \ll N.$$

A linear combination of elements of $X(k)$, given by

$$y(m) = \sum_{k=0}^{N-1} a_m(k) X(k), \tag{1}$$

is called a measurement, with the weights denoted by $a_m(k)$.

The above set of measurements, $y(m), m = 0, 1, \ldots, M-1$, admits a vector/matrix form, given by

$$\mathbf{y} = \mathbf{A}\mathbf{X}, \tag{2}$$

where $\mathbf{y} = [y(0), y(1), \ldots, y(M-1)]^T$ is an $M \times 1$ column vector, $\mathbf{A}$ is an $M \times N$ measurement matrix which comprises the weights $a_m(k)$ as its elements, and $\mathbf{X}$ is an $N \times 1$ sparse column vector with elements $X(k)$.

Without loss of generality, we shall assume that the measurement matrix, $\mathbf{A}$, is normalized so that the energy of its columns sums up to unity. Consequently, the diagonal elements of its symmetric Gram form, $\mathbf{A}^H\mathbf{A}$, are equal to 1, where $\mathbf{A}^H$ is the complex conjugate transpose of $\mathbf{A}$.

The compressive sensing theory task is to reconstruct the $N$-dimensional $K$-sparse vector $\mathbf{X}$ from a set of $M$ measurements, $\mathbf{y} = \mathbf{A}\mathbf{X}$, with $K \ll M < N$. There are several approaches to solve this problem (for reviews of these approaches see [3], [13]). Here we will consider the orthogonal matching pursuit (OMP) approach [3], [5], [6], [13].

### B. OMP Solution to the CS Paradigm

A matching pursuit reconstruction algorithm is typically based on a two-step strategy:

Manuscript received March 11, 2021; revised May 15, 2021; accepted May 19, 2021. Date of publication May 28, 2021; date of current version June 11, 2021. This work was supported by the Montenegrin Academy of Sciences and Arts (CANU). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Yipeng Liu. *(Corresponding author: Ljubisa Stankovic.)*

Ljubiša Stanković, Miloš Brajović, Isidora Stanković, and Miloš Daković are with the University of Montenegro, Podgorica 81000, Montenegro (e-mail: ljubisa@ac.me; milosb@ac.me; isidoras@ucg.ac.me; milos@ac.me).

Danilo Mandic is with the Imperial College London, London SW7 2BT, U.K. (e-mail: d.mandic@imperial.ac.uk).

Digital Object Identifier 10.1109/LSP.2021.3084559

Step 1: Detect the positions of nonzero elements,
Step 2: Recover the signal by exploiting the relations between the measurement matrix, $\mathbf{A}$, detected positions, and the vector of measurements, $\mathbf{y}$.

It will be further shown that the physically relevant conditions for the reconstruction are in fact related to the challenges emerging in the first step of the presented methodology. Otherwise, if arbitrary positions of $K$ nonzero elements of $\mathbf{X}$ are known, meaning that $X(k) \neq 0$ for $k \in \{k_1, k_2, \ldots, k_K\}$, then a reduced set of measurement equations will follow as

$$\mathbf{y} = \mathbf{A}_{MK}\mathbf{X}_K,$$

with $\mathbf{A}_{MK}$ being an $M \times K$ dimensional sub-matrix of the matrix $\mathbf{A}$, formed by keeping only the columns corresponding to the positions $\{k_1, k_2, \ldots, k_K\}$. Unknown values $X(k)$, located at $k \in \{k_1, k_2, \ldots, k_K\}$, are here conveniently grouped into a $K \times 1$ vector $\mathbf{X}_K$. This system of $M$ equations and $K < M$ unknowns has a solution in a Least Square (LS) sense,

$$\mathbf{X}_K = (\mathbf{A}_{MK}^H\mathbf{A}_{MK})^{-1}\mathbf{A}_{MK}^H\mathbf{y} = \text{pinv}(\mathbf{A}_{MK})\mathbf{y}. \quad (3)$$

A sufficient condition for this reconstruction with known positions is that the matrix $\mathbf{A}_{MK}^H\mathbf{A}_{MK}$ is regular.

*The reconstruction solution is exact if the positions $\{k_1, k_2, \ldots, k_K\}$ of nonzero elements in a $K$-sparse vector $\mathbf{X}$ are exactly determined for any set $\{k_1, k_2, \ldots, k_K\}$ and if there exist at least $K$ independent measurements [5]–[8].*

The more demanding condition is that the positions of the nonzero elements in the sparse vector are exactly determined. This means that the detection step of the OMP approach is crucial for the exact solution. The detection is based on the initial estimate, defined as a back-projection of the measurements, $\mathbf{y}$, on the measurement matrix, $\mathbf{A}$, given by

$$\mathbf{X}_0 = \mathbf{A}^H\mathbf{y} = (\mathbf{A}^H\mathbf{A})\mathbf{X}. \quad (4)$$

If $\mathbf{A}^H\mathbf{A}$ ensures that the largest $K$ elements of the initial estimate, $\mathbf{X}_0$, are positioned at exact $k \in \{k_1, k_2, \ldots, k_K\}$, then the detection is performed by taking the positions of the highest magnitude elements in the initial estimate, followed by the reconstruction based on (3). The condition that $K$ elements in the initial estimate, $\mathbf{X}_0$, located at the positions of non-zero elements in the original sparse vector, $\mathbf{X}$ are larger than any other component in the initial estimate can be relaxed through an iterative procedure. In such a methodology [5], in order to find the position $k_1$ of the largest non-zero element in $\mathbf{X}_0$, it is required that its value is larger than any value at the original zero-valued element position. Upon detecting the position, $k_1$, and estimating the component value based on (3), with $\mathbf{A}_K$ being formed based on $k_1$, the contribution of this component is removed from measurements vector, $\mathbf{y}$. The procedure is iteratively repeated for the remaining nonzero elements.

## II. Uniqueness of the OMP Reconstruction

The uniqueness condition based on the coherence index can be formulated as follows: *The reconstruction of a $K$-sparse signal, $\mathbf{X}$, is unique if the coherence index,*

$$\mu = \max_{k,l\ k\neq l}\left|\sum_{m=0}^{M-1} a_m(k)a_m^*(l)\right|, \quad (5)$$

*of the normalized measurement matrix, $\mathbf{A}$, satisfies [2]*

$$K < \frac{1}{2}\left(1 + \frac{1}{\mu}\right). \quad (6)$$

The coherence index, $\mu$, is equal to the maximum absolute off-diagonal element of $\mathbf{A}^H\mathbf{A}$, while its diagonal elements are equal to 1. The condition in (6) guarantees that the solutions obtained by minimizing the $\ell_0$-norm and $\ell_1$-norm produce the same common unique solution [12]. This condition guarantees uniqueness of the solution produced by the OMP algorithm [4]–[10].

The coherence index condition (5) can be obtained through the analysis of successful determination of the positions of non-zero values in the original vector $\mathbf{X}$ [10]. By definition, any measurement represents a linear combination of nonzero elements of the sparse vector $\mathbf{X}$, that is

$$y(m) = \sum_{i=1}^{K} X(k_i)a_m(k_i).$$

Furthermore, without loss of generality, it can be assumed that the largest element is $X(k_1) = 1$, whereas the remaining nonzero elements do not take values greater than this value, $|X(k_i)| \leq 1$, $i = 2, 3, \ldots, K$. In that case, the initial estimate

$$X_0(k) = \sum_{m=0}^{M-1}\left[\sum_{i=1}^{K} X(k_i)a_m(k_i)\right]a_m^*(k) = \sum_{i=1}^{K} X(k_i)\mu(k_i, k)$$

can be expressed, for the element at $k = k_1$, as follows

$$X_0(k_1) = X(k_1) + \sum_{i=2}^{K} X(k_i)\mu(k_i, k_1), \quad (7)$$

where $\mu(k_i, k) = \sum_{m=0}^{M-1} a_m(k_i)a_m^*(k)$. The maximum possible absolute value of $\mu(k_i, k)$ is then equal to the coherence index, that is, $\mu = \max_{k_i,k} |\mu(k_i, k)|$, $k \neq k_i$.

In the worst case scenario for the detection of the element at position $k_1$, the value of this element, $|X_0(k_1)|$ in (7) is maximally reduced by the term $\sum_{i=2}^{K} X(k_i)\mu(k_i, k)$. The maximally reduced coefficient $|X_0(k_1)|$ takes the value

$$\min|X_0(k_1)| = 1 - \sum_{i=2}^{K}|X(k_i)\mu(k_i, k)| = 1 - (K-1)\mu, \quad (8)$$

assuming that all $K-1$ remaining elements $X(k_i)$ have the most unfavorable value, $X(k_i) = 1$, whereas $|\mu(k_i, k)| = \mu$, for each $k_i \in \{k_1, k_2, \ldots, k_K\}$. The maximum value of disturbance at the positions where the elements were originally zero-valued, $k \notin \{k_1, k_2, \ldots, k_K\} = \mathbb{K}$ is equal to

$$\max_{k,k\notin\mathbb{K}}\left|\sum_{i=1}^{K} X(k_i)\mu(k_i, k)\right| = K\mu. \quad (9)$$

In the worst case scenario, the exact and unique detection of the position of the largest element $X_0(k_1)$ is possible when its maximally degraded value exceeds the maximal value of the disturbance

$$\min|X_0(k_1)| > \max_{k,k\notin\mathbb{K}}\left|\sum_{i=1}^{K} X(k_i)\mu(k_i, k)\right|,$$

or equivalently, $1 - (K-1)\mu > K\mu$, producing (6). Upon successfully detecting, reconstructing, and removing the first nonzero component in a sparse $\mathbf{X}$, the same procedure and relations

can be iteratively applied to the remaining "deflated" signal, whose sparsity cannot be greater than $K$, [7], guaranteeing an exact and unique solution as shown in [7], [8].

## III. Improved Bound Derivation

In the previous derivation of the reconstruction relation (6), it has been assumed that $K$ maximum absolute values of $\mu(k_i,k)=\mu$ in (9) add up to form the disturbance. Moreover, it has been assumed that the component $X_0(k_1)$, that we aim to detect at a position $k_1$, is reduced by $K-1$ maximal values of $\mu(k_i,k)=\mu$. This is, however, an overly pessimistic assumption, since even in the worst case scenario all the largest $2K-1$ values of $\mu(k_i,k)$, in general, may not be equal to $\mu$.

Actually, when the first maximum is taken $|\mu(k_i,k)|=\mu$, in the next sample only the second largest value of $|\mu(k_i,k)|$ can be taken. Subsequently, only the third largest value of $|\mu(k_i,k)|$ can be taken, and so on. To take this fact into account and derive a less conservative reconstruction bound, denote the sorted values of $|\mu(k_i,k)|$ by

$$s(p)=\text{sort}_{k_i,k}\{|\mu(k_i,k)|\},$$
$$k_i,k=0,1,\ldots,N-1,k_i\neq k,\;\; p=1,2,\ldots,N(N-1), \tag{10}$$

with a nonincreasing order, $s(1)\geq s(2)\geq\cdots\geq s(N^2-N)$. In the worst case scenario, instead of $2K-1$ values of $\mu$, now we can use the first $(2K-1)$ (largest) values of $s(p)$ to get

$$1>\sum_{p=1}^{2K-1}s(p)=(2K-1)\alpha_{\mathbf{A}}(2K-1),$$

instead of $1>(K-1)\mu+K\mu$, where $\alpha_{\mathbf{A}}(2K-1)$ is the mean value of the $(2K-1)$ largest values of $|\mu(k_i,k)|$,

$$\alpha_{\mathbf{A}}(2K-1)=\frac{1}{2K-1}\sum_{p=1}^{2K-1}s(p)=\underset{1\leq p\leq 2K-1}{\text{mean}}\,s(p).$$

The bound for the reconstruction now becomes

$$K<\frac{1}{2}\left(1+\frac{1}{\alpha_{\mathbf{A}}(2K-1)}\right). \tag{11}$$

This implicit inequality is easily solved by direct check, starting from $K=1$, followed by increasing the value of $K$ by one, until the inequality (11) holds [10].

In the special case of an equiangular tight frame (ETF) measurement matrix, when the factor $|\mu(k_i,k)|=\mu$ is constant, then $\alpha_{\mathbf{A}}(2K-1)=\mu$ and (6) holds. In all cases, for any measurement matrix $\mathbf{A}$, condition $\alpha_{\mathbf{A}}(2K-1)\leq\mu$ holds. This means that a more optimistic bound for $K$ is obtained by (11) than the conventional CS bound in (6).

Furthermore, it will be shown that a less conservative bound than in (11) can be derived following some simple observations of the initial estimate calculation based on the Gram matrix $\mathbf{A}^H\mathbf{A}$. Recall that the value of the initial estimate, $X_0(k)$, at a non-zero position $k_1$ can be calculated using (7). In the described worst case scenario, the observed "largest" term $X(k_1)=1$ is maximally reduced. This happens when $|\mu(k_1,k_i)|$ takes the largest possible values *only within a row with index* $k_1$ of matrix $\mathbf{A}^H\mathbf{A}$. If we sort the values of elements within each row in a nonincreasing order (for all indices $k_1=k$, since $k_1$ can assume any index value) and form

$$s_1(k,p)=\text{sort}|\mu(k,l)| \tag{12}$$

$l=0,1,\ldots,N-1,\;\; l\neq k,\;\; p=1,2,\ldots,N-1,$ such that $s_1(k,1)\geq s_1(k,2)\geq\cdots\geq s_1(k,N-1)$, then in the worst case scenario, when $X(k)=1$ for all $k\in\{k_1,k_2,\ldots,k_K\}$, $X(k_1)$ in (7) will be reduced for

$$(K-1)\max_{k_1}\left\{\frac{1}{K-1}\sum_{p=1}^{K-1}s_1(k_1,p)\right\}=(K-1)\beta_{\mathbf{A}}(K-1).$$

That is, $X(k_1)$ is reduced by the sum of the first $(K-1)$ terms of $s_1(k_1,p)$, being in fact the largest possible $(K-1)$ values of any row in $|\mu(k_1,l)|$, $k_1=0,1,\ldots,N-1$, excluding the element $l=k_1$, where

$$\beta_{\mathbf{A}}(K-1)=\max_{k_1}\{\underset{1\leq p\leq K-1}{\text{mean}}\,s_1(k_1,p)\}. \tag{13}$$

The largest possible disturbance value is obtained when absolute values of the elements $|\mu(k_i,k)|$ of matrix $\mathbf{A}^H\mathbf{A}$ take the largest values *in the given row at the disturbance position* $k\notin\{k_1,k_2,\ldots,k_K\}$ and are summed in phase. If we take into account the notation for sorted values in (12), then this accumulated disturbance value becomes $K\gamma_{\mathbf{A}}(K)$, where

$$\gamma_{\mathbf{A}}(K)=\max_{k\neq k_1}\{\underset{1\leq p\leq K}{\text{mean}}\,s_1(k,p)\} \tag{14}$$

having in mind that $|\mu(k_i,k)|$ can take $K$ largest values from *one row* (excluding the values in the row $k_1$ taken in (13)).

The successful detection of the component $X(k_1)$ will not be compromised if this component, assuming its smallest possible value, is still larger than the largest value of the disturbance at $k\notin\{k_1,k_2,\ldots,k_K\}$, that is

$$1-\sum_{i=2}^{K}|X(k_i)\mu(k_i,k)|>\sum_{i=1}^{K}|X(k_i)\mu(k_i,k)|$$

or, having in mind (13) and (14)

$$1>(K-1)\beta_{\mathbf{A}}(K-1)+K\gamma_{\mathbf{A}}(K).$$

*The reconstruction of a K-sparse signal,* $\mathbf{X}$*, is exact and unique if the measurement matrix,* $\mathbf{A}$*, guarantees that the following condition is satisfied*

$$K<\frac{1+\beta_{\mathbf{A}}(K-1)}{\beta_{\mathbf{A}}(K-1)+\gamma_{\mathbf{A}}(K)}. \tag{15}$$

The three discussed sparsity bounds are related as

$$\frac{1+\beta_{\mathbf{A}}(K-1)}{\beta_{\mathbf{A}}(K-1)+\gamma_{\mathbf{A}}(K)}\geq\frac{1}{2}\left(1+\frac{1}{\alpha_{\mathbf{A}}(2K-1)}\right)\geq\frac{1}{2}\left(1+\frac{1}{\mu}\right).$$

The equality holds for the ETF measurement matrices when $\beta_{\mathbf{A}}(K-1)=\gamma_{\mathbf{A}}(K)=\alpha_{\mathbf{A}}(2K-1)=\mu$.

## IV. Numerical Examples

The presented relations are tested on several numerical examples: a graph matrix, a measurement matrix of a Gaussian form, partial DFT and DCT matrices, and an ETF form.

An unweighted and undirected graph with $N=64$ vertices is given in Fig. 1(left). The graph Fourier transform (GFT) matrix is defined by the eigenvectors of the graph Laplacian, as its columns [16]. It has been assumed that the graph signal is $K$-sparse in the GFT domain and that the samples at vertices $n=21$ and $n=38$ are missing. The off-diagonal absolute values of $\mathbf{A}^H\mathbf{A}$, with $\mathbf{A}$ corresponding to $M=62$ available

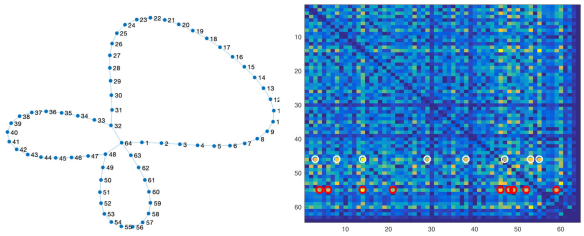

Fig. 1. An undirected and unweighted graph (left) and the absolute values of matrix $\mathbf{A}^H\mathbf{A}-\mathbf{I}$ (right), where $\mathbf{A}$ is obtained from the graph eigenvector matrix, excluding rows 21 and 38. The elements of matrix $\mathbf{A}^H\mathbf{A}$ used for the calculation of $\beta_{\mathbf{A}}(K-1)$ are encircled using a white line, while its elements used for the calculation of $\gamma_{\mathbf{A}}(K)$ are encircled using a red line, with sparsity $K=9$, being the smallest $K$ when inequality (15) is not satisfied.

samples (rows of the GFT matrix), is shown in Fig. 1(right). This matrix guarantees unique reconstruction for $K<6.8917$, $K<7.4618$, and $K<8.3118$, respectively, with the three presented approaches for the sparsity bound, given by relations (6), (11), and (15), respectively. We can conclude that the sparsity limit is improved from the maximum sparsity $K=6$, with (5), to $K=8$ using (15). The results are statistically checked. It has been concluded that in $10^6$ random realizations for $K\leq 8$ all reconstructions were successful.

For a Gaussian measurement matrix, we used $N=80$ and $M=70$. Over 1000 realizations (with various Gaussian matrices) we obtained the mean value of the limit $K<1.6761$ with a standard deviation (SD) of 0.08, while the presented method produced the mean value limit $K<2.3523$ with an SD of 0.16. The same experiment with a $900\times 1000$ measurement matrix produced $K<3.6175$ with an SD of 0.15 and $K<4.3580$ with an SD of 0.07, meaning that the sparsity of a certain reconstruction is improved from 3 to 4.

For a partial DCT matrix, with dimension $124\times 128$, we obtained the mean values $K<9.7849$ and $K<12.1354$, with (6) and (15). The best case in 1000 random realizations of the available samples, with improved bound, was $K<14.2238$.

For an $124\times 128$ partial DFT matrix, we obtained the mean values $K<16.9068$ and $K<19.8323$, with (6) and (15), respectively. The best case over 1000 random realizations of the available samples was $K<21.4307$ with (15). For just 2 missing samples, the measurement matrix behaves close to an ETF and produced the limits for $K$ very close to $N/4=32$. For $112\times 128$ measurement matrix the best limit was $K<10.3770$. For half of the available samples, $M=N/2$, all limits dropped toward the theoretically worst case when no unique solution can be achieved. For 1000 realizations, the mean values were just above 3, the best form of the measurement matrix produced the sparsity limit slightly above 4, while for the worst case measurement matrix in these 1000 realizations, the limit dropped to $K<1.9545$. With a $20\times 128$, corresponding to just 20 available samples, the mean limits were $K<1.6325$, and $K<2.2649$.

Finally, for an ETF matrix of dimension $9\times 18$, all the presented limits were the same, as expected since the absolute values of the off-diagonal elements of $\mathbf{A}^H\mathbf{A}$ are the same. The common limit was $K<2.5616$.

*Remarks:* (i) Notice that (6) is the worst case recovery condition. A successful recovery could still be achieved even if this condition is not satisfied [10]. (ii) In the case when (6) is not met, the reconstruction can be improved if the algorithm is allowed to run more than $K$ iterations [6], [10]. (iii) Additional improvement of the bound in (15) can be achieved if all indices used in (13) are considered as the positions of nonzero elements, and excluded in (14), instead of excluding $k_1$ only. (iv) Probabilistic analysis may be found in [10], [17], [18]. (v) The condition for a unique solution existence can be relaxed using (16) or [19].

## V. Generalization for Two Bases and $\ell_0$-Norm

The presented framework can be used to generalize the results obtained analyzing the signal representation in two bases, as introduced in [14]. This kind of signal representation was used to find the general sparsity bounds for the unique solutions, obtained using the $\ell_0$-norm and the $\ell_1$-norm minimizations. Consider representations of a unit energy signal, $x(n)$, $n=0,1,\ldots,N-1$, in two arbitrary bases $u_k(n)$ and $v_l(n)$, with the respective transformation elements $X(k)$ and $Y(l)$, $k,l=0,1,\ldots,N-1$. Assume, as in [2], [14], that the signal is sparse in these bases with sparsities $||\mathbf{X}||_0=K$ and $||\mathbf{Y}||_0=L$, and that the Parseval's theorem holds in both bases, $||\mathbf{X}||_2^2=1$ and $||\mathbf{Y}||_2^2=1$. We can now form a function $L(n,k,l)=X(k)Y^*(l)u_k(n)v_l^*(n)$, as in [15], such that

$$\sum_n x(n)x^*(n)=\sum_n\sum_k\sum_l X(k)Y^*(l)u_k(n)v_l^*(n)=1,$$

where $k\in\{k_1,k_2,\ldots,k_K\}$ and $l\in\{l_1,l_2,\ldots,l_L\}$. Then, using Schwartz's inequality, with $\mu(k,l)=\sum_n u_k(n)v_l^*(n)$, we get

$$\begin{aligned}1&=\Big|\sum_k\sum_l X(k)Y^*(l)\sum_n u_k(n)v_l^*(n)\Big|^2\\&\leq\sum_k\sum_l|X(k)|^2|Y(l)|^2\sum_k\sum_l|\mu(k,l)|^2\\&\leq KL\frac{1}{KL}\sum_p s^2(p)\end{aligned}$$

where $s(p)$ is defined in (10). Using the notation $\eta_{\mathbf{A}}(KL)=\frac{1}{KL}\sum_p s^2(p)$ and $\sqrt{ab}\leq(a+b)/2$, $a>0$, $b>0$, we get

$$\frac{1}{\eta_{\mathbf{A}}(KL)}\leq KL=||\mathbf{X}||_0||\mathbf{Y}||_0\leq\left(\frac{1}{2}(||\mathbf{X}||_0+||\mathbf{Y}||_0)\right)^2$$

$$\text{or } ||\mathbf{X}||_0+||\mathbf{Y}||_0\geq\frac{2}{\sqrt{\eta_{\mathbf{A}}(KL)}}.$$

The solution of the $\ell_0$-norm minimization is unique if the sparsity, $||\mathbf{X}||_0$, is smaller than half of the uncertainty bound

$$K<\frac{1}{\sqrt{\eta_{\mathbf{A}}(K^2)}},\tag{16}$$

for $K=L$. This relation can be used to improve coherence index-based conditions when the $\ell_0$-norm and $\ell_1$-norm minimization produce the same and unique solution [2], [14].

## VI. Conclusion

A numerically efficient calculation of an improved coherence index-based sparsity bound has been proposed. The calculation is demonstrated on a graph signal example and over several commonly used measurement matrices. The results are generalized for the $l_0$-norm and two bases.

## REFERENCES


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

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

[3] L. Stanković, E. Sejdić, S. Stanković, M. Daković, and I. Orović. "A tutorial on sparse signal reconstruction and its applications in signal processing." *Circuits, Syst., Signal Process.*, vol. 38, no. 3, pp. 1206–1263, Mar. 2019.

[4] D. L. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, Jan. 2006.

[5] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[6] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[7] T. Cai, L. Wang and G. Xu, "Stable recovery of sparse signals and an oracle inequality," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3516–3522, Jul. 2010.

[8] J. Wang and B. Shim, "On the recovery limit of sparse signals using orthogonal matching pursuit," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4973–4976, Sep. 2012.

[9] R. S. Varga, *Geršgorin and His Circles*. Berlin, Germany: Springer-Verlag, 2004.

[10] L. Stanković, D. Mandic, M. Daković, and I. Kisil, "Demystifying the coherence index in compressive sensing," *IEEE Signal Process. Mag.*, vol. 37, no. 1, pp. 152-162, Jan. 2020.

[11] 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.

[12] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.

[13] M. Rani, S. B. Dhok, and R. B. Deshmukh, "A systematic review of compressive sensing: Concepts, implementations and applications," *IEEE Access*, vol. 6, pp. 4875–4894, 2018.

[14] M. Elad and A. M. Bruckstein, "A. generalized uncertainty principle and sparse representation in pairs of bases," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2558–2567, Sep. 2002.

[15] L. Stanković, "The support uncertainty principle and the graph Rihaczek distribution: Revisited and improved," *IEEE Signal Process. Lett.*, vol. 27, pp. 1030–1034, 2020.

[16] L. Stanković *et al.*, "Data analytics on graphs part I: Graphs and spectra on graphs," *Found. Trends Machine Learn.*, vol. 13, no. 1, pp. 1–157, 2020.

[17] I. Stankovic, M. Brajovic, M. Dakovic, C. Ioana, and L. Stankovic, "Quantization in compressive sensing: A signal processing approach," in *IEEE Access*, vol. 8, pp. 50611–50625, 2020, doi: 10.1109/ACCESS.2020.2979935.

[18] L. Stankovic, and M. Brajovic, "Analysis of the reconstruction of sparse signals in the DCT domain applied to audio signals," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no.7, pp. 1216–1231, Jul. 2018.

[19] L. Stankovic, "On the sparsity bound for the existence of a unique solution in compressive sensing by the gershgorin theorem," *arXiv:2105.13198*, 2021.