Ljubiša Stanković, Danilo P. Mandic, Miloš Daković, and Ilya Kisil

Demystifying the Coherence Index in Compressive Sensing

The existence and uniqueness conditions are a prerequisite to ensure the reliable reconstruction of sparse signals from reduced sets of measurements within the compressive sensing (CS) paradigm. However, despite their underpinning role in practical applications, the existing uniqueness relations are either computationally prohibitive to implement [the restricted isometry property (RIP)] or involve mathematical tools that are beyond the standard background of engineering graduates (the coherence index). This may introduce conceptual and computational obstacles in the development of engineering intuition, design of suboptimal practical solutions, and understanding of theoretical and practical limitations of the CS framework.

To this end, we employ standard linear algebra to introduce a simple but rigorous derivation of the coherence index condition, with the aim of empowering signal processing practitioners with intuition for the design and implementation of CS systems. Given that the coherence index is one of very few CS metrics that admits mathematically tractable and computationally feasible calculation, we hope that this article will help to bridge the gap between the theory and applications of CS.

The basic CS setting

CS is a maturing field that, under appropriate conditions, provides a rigorous framework for efficient data acquisition

[1], [2]. 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.

Research under an umbrella of sparsity has been a major topic of investigation for about a quarter of century and produced solid theory to support the exact reconstruction of sparse signals in CS scenarios. The coherence index is of particular interest for practitioners since it is one of the very few supporting theoretical tools within the CS framework that can be calculated in a computationally feasible way. However, its derivation follows a rather complex and convoluted path; this fact spurred us to revisit the coherence index from a signal processing perspective to equip practitioners with an easy-to-understand interpretation of and intuition for the design and physical meaningfulness in the analysis. We provide a similar intuition for the feasibility of calculating the RIP condition [1]–[9].

Definitions and notation

Definition 1

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

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

Definition 2

A linear combination of the elements of $X(k)$ that is given by

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

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

The sensing scheme for producing the measurements $y(m)$, $m = 0, 1, \ldots, M - 1$, admits a vector/matrix form that is given by

$$\mathbf{y} = \mathbf{A}\mathbf{X}, \qquad (2)$$

where $\mathbf{y} = [y(0), y(1), \ldots, y(M-1)]^T$ is an $M \times 1$ column vector of the measurements, $\mathbf{A}$ is an $M \times N$ measurement matrix that has the weights $a_m(k)$ as its elements, and $\mathbf{X}$ is an $N \times 1$ sparse column vector with elements $X(k)$. An illustration of the CS concept is given in Figure 1.

Without a 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 one, where $\mathbf{A}^H$ is the complex conjugate transpose of $\mathbf{A}$.

While CS theory states that, under certain mild conditions, it is possible to reconstruct a sparse N-dimensional vector, $\mathbf{X}$, from a reduced M-dimensional set of measurements, $\mathbf{y}$, practitioners require physically meaningful, intuitive, and easily interpretable uniqueness

tools for the applications of CS to become more widespread.

A solution to the CS paradigm

Several approaches have been established for the CS paradigm, and we follow the principles behind the matching-pursuit method. The simplest case, when the positions of the nonzero elements in $\mathbf{X}$ are known, is considered first to provide the intuition and an example of the detection of unknown positions of the nonzero elements in $\mathbf{X}$. This serves as a basis for our derivation of the uniqueness relation through the coherence index for a general case of unknown positions of the nonzero elements in $\mathbf{X}$.

Known coefficient positions

Consider the case with K nonzero elements of $\mathbf{X}$ located at arbitrary but known positions; that is, $X(k) \neq 0$ for $k \in \{k_1, k_2, …, k_K\}$. Compared to the general form of the measurement relations in (2), this gives rise to the following reduced system of equations:

$$
\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(M-1) \end{bmatrix} = \begin{bmatrix} a_0(k_1) & \dots & a_0(k_K) \\ a_1(k_1) & \dots & a_1(k_K) \\ \vdots & \ddots & \vdots \\ a_{M-1}(k_1) & \dots & a_{M-1}(k_K) \end{bmatrix}
$$
$$
\times \begin{bmatrix} X(k_1) \\ X(k_2) \\ \vdots \\ X(k_K) \end{bmatrix}. \tag{3}
$$

For a successful CS recovery, the matrix form of the system, which is given by

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

needs to be solved for the nonzero elements $X(k)$ located at $k \in \{k_1, k_2, …, k_K\}$, which are conveniently grouped into a $K \times 1$ vector, $\mathbf{X}_K$. Observe that the matrix, $\mathbf{A}_{MK}$, is an $M \times K$ dimensional submatrix of the full $M \times N$ measurement matrix, $\mathbf{A}$, in (2), whereby only the columns that correspond to the positions of the nonzero elements in $\mathbf{X}$ are kept (for illustration, see columns 1 and 6 in Figure 1). The smallest number of measurements needed to

recover the K-element vector, $\mathbf{X}_K$, is, therefore, $M = K < N$. For $M > K$, as in Figure 1, the system is overdetermined, and the solution is found in the least-squares sense to give [3]

$$\mathbf{X}_K = (\mathbf{A}_{MK}^H \mathbf{A}_{MK})^{-1} \mathbf{A}_{MK}^H \mathbf{y} = \text{pinv}(\mathbf{A}_{MK})\mathbf{y}, \tag{4}$$

where $\text{pinv}(\mathbf{A}_{MK}) = (\mathbf{A}_{MK}^H \mathbf{A}_{MK})^{-1} \mathbf{A}_{MK}^H$ denotes the pseudo-inverse of matrix $\mathbf{A}_{MK}$, while the matrix $\mathbf{A}_{MK}^H \mathbf{A}_{MK}$ is referred to as the $K \times K$ dimensional *Gram matrix* of $\mathbf{A}_{MK}$.

From (4), the existence of a recovery solution requires that the inverse $(\mathbf{A}_{MK}^H \mathbf{A}_{MK})^{-1}$ does exist or, in other words, that $\text{rank}(\mathbf{A}_{MK}^H \mathbf{A}_{MK}) = K$. This stipulation can be equally expressed via the matrix condition number as $\text{cond}(\mathbf{A}_{MK}^H \mathbf{A}_{MK}) < \infty$, which casts the existence condition into a more convenient form that is dictated by the eigenvalue spread of $\mathbf{A}_{MK}^H \mathbf{A}_{MK}$. For noisy measurements, the reconstruction error includes contributions from the input noise and ill-posedness that are due to a high matrix-condition number (known as the *mathematical artifact*). In this sense, the goal of a successful CS (sampling) strategy can be interpreted as that of forming the measurement matrix, $\mathbf{A}$, to produce a condition number as close to 1 as possible.

The signal processing framework for CS

The most fertile domains for signal sparsity are common linear signal transforms, since they enable the original time-domain samples of the signal $x(n)$ to be considered as measurements (linear combinations) of the representation domain coefficients, $X(k)$. For example, when the discrete Fourier transform (DFT) is used as the signal sparsity domain, the measured signal samples can be expressed as

$$y(m) = x(n_m) = \frac{1}{\sqrt{M}} \sum_{k=0}^{N-1} e^{j2\pi n_m k/N} X(k) \tag{5}$$

for $m = 0, 1, …, M-1$. In this case, the measurements, $y(m)$, can be regarded as a reduced set of signal samples, $x(n_m)$. Observe that this expression conforms to the general CS formulation (1), where the weights $a_m(k) = \exp(j2\pi n_m k/N)/\sqrt{M}$. Since CS employs a random subset of time instants, $\{n_0, n_1, …, n_{M-1}\} \subset \{0, 1, 2, …, N-1\}$, measurement matrix $\mathbf{A}$ is obtained from the inverse DFT matrix, whereby only the rows corresponding to $\{n_0, n_1, …, n_{M-1}\}$ are kept. Matrix $\mathbf{A}$, obtained in such a way, is called a *partial DFT measurement matrix*.

Remark 1

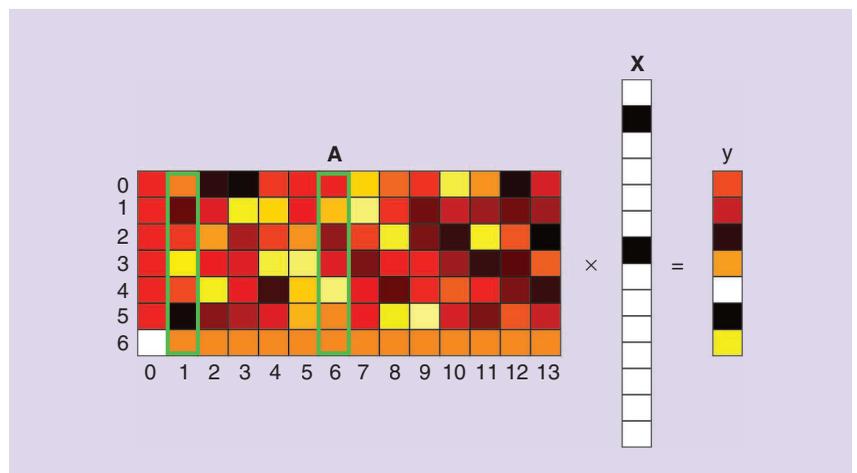The convenience of the considered DFT representation enables us to consider

FIGURE 1. The CS principle. The short and wide measurement matrix, **A**, maps the original *N*-dimensional *K*-sparse vector, **X**, to an *M*-dimensional dense vector of measurements, **y**, with $M < N$ and $K \ll N$. In our case, $N = 14$, $M = 7$, and $K = 2$. Since M is the maximum rank of **A** and $M < N$, the original *N*-dimensional vector **X** cannot, in general, be recovered from the measurements, **y** (there is lack of degrees of freedom). However, the CS paradigm allows for the complete and unique recovery of sparse signals, with the coherence index being a common way to define the corresponding conditions. The same **A** is used in Figure 2.

Nyquist subsampling as a special case of a reduced set of measurements, with the specific positions of K nonzero elements given by

$$\mathbf{X} = [X(0), X(1), \ldots, X(K-1), \\ 0, 0, \ldots, 0]^T.$$

The classic subsampling operation can be explained by starting from the assumption that only the first $K = N/P$ elements are nonzero, with $P > 1$ an integer. Then, the original signal can be subsampled at $n_m = mP$, since this yields $a_m(k) = \exp(j2\pi mk/K)/\sqrt{K}$, with $\mathbf{A}_{MK}^H \mathbf{A}_{MK}$ being an identity matrix. On the other hand, some important applications, such as radar signal processing, routinely deal with a very small number of nonzero elements, $X(k)$. However, in contrast to the classic subsampling scenario, these nonzero transform-domain elements may be located at any position $k \in \{0, 1, \ldots, N-1\}$. While this makes it impossible to perform classical subsampling, this class of applications does admit a unique solution with a reduced number of signal samples within the CS theory framework.

Despite the obvious methodological advantages of CS over classical analyses, operation on a reduced set of measurements compromises the uniqueness of the CS solution. It is, therefore, natural to first examine the conditions that should be satisfied by the measurements (their number and the properties of the measurement matrix) so that the existence of the CS solution is guaranteed and unique.

Example 1
To illustrate how a reduced set of samples can compromise the uniqueness of the solution, consider a signal with the total length of N and the smallest sparsity degree of $K = 1$ in the DFT domain, that is, with only one nonzero element in $\mathbf{X}$. Assume that a significant number of $M = N/2$ measurements (signal samples) are available at $n = 0, 2, \ldots, N-4, N-2$, with an even N, and suppose that the measurement values are $y(m) = x(2m) = 1$. The solution to this simple problem is not unique, since $x(n) = \exp(j2\pi nk_1/N)$, $n = 0, 1, \ldots, N-1$, for $k_1 = 0$ and $k_1 = N/2$, satisfying all of the problem conditions.

The uniqueness of the CS paradigm
In general, the question of uniqueness can be understood within the following framework. Consider a K-sparse vector $\mathbf{X}$ with the nonzero elements $X(k) \neq 0$ at $k \in \{k_1, k_2, \ldots, k_K\}$ and assume that its vector form, $\mathbf{X}_K$, is a solution to $\mathbf{y} = \mathbf{A}_{MK} \mathbf{X}_K$. Assume also that the solution is not unique, so there exists another vector, $\mathbf{X}'$, with nonzero elements at the different positions $k \in \{k_{K+1}, k_{K+2}, \ldots, k_{2K}\}$, whose reduced form, $\mathbf{X}'_K$, supports the same measurements, $\mathbf{y}$; that is, $\mathbf{y} = \mathbf{A}'_{MK} \mathbf{X}'_K$. Then, $\mathbf{A}_{MK} \mathbf{X}_K - \mathbf{A}'_{MK} \mathbf{X}'_K = \mathbf{0}$, and this matrix equation can be combined into $\mathbf{A}_{M2K} \mathbf{X}_{2K} = \mathbf{0}$, where $\mathbf{A}_{M2K}$ is an $M \times 2K$-dimensional submatrix of the measurement matrix, $\mathbf{A}$, and $\mathbf{X}_{2K}$ is a $2K$-dimensional vector. Then,

- A nontrivial solution to the matrix equation $\mathbf{A}_{M2K} \mathbf{X}_{2K} = \mathbf{0}$ indicates that the CS solution is nonunique. The condition for a nonunique solution is, therefore, $\text{rank}(\mathbf{A}_{M2K}^H \mathbf{A}_{M2K}) < 2K$ for at least one combination of $2K$ nonzero element positions.
- If $\text{rank}(\mathbf{A}_{M2K}^H \mathbf{A}_{M2K}) = 2K$ for all possible combinations of $2K$ nonzero element positions (out of N), the scenario of two K-sparse solutions is not possible, and the solution is unique.

This rationale is a starting point for the definition of the common uniqueness criteria in CS. The coherence-index relation is typically derived through the Gershgorin disk theorem [10]. Notice that the maximum robustness of the condition $\text{rank}(\mathbf{A}_{M2K}^H \mathbf{A}_{M2K}) = 2K$ is achieved if $\text{cond}(\mathbf{A}_{MK}^H \mathbf{A}_{MK})$ is close to unity. It should be mentioned that the RIP is tested in a similar way through a combinatorial consideration of the matrix norm $\|\mathbf{A}_{M2K} \mathbf{X}_{2K}\|_2^2 / \|\mathbf{X}_{2K}\|_2^2$ through eigenvalue analysis. The general case, with unknown positions of the nonzero elements in $\mathbf{X}$, can be solved by combining the method for the known positions and a direct search approach, as described in "The Direct Search, Uniqueness of the Solution, and Restricted Isometry Property." However, this is not computationally feasible.

Structured and sparse measurement matrices
It is common for the measurement matrices to combine all of the elements from $\mathbf{X}$ into every measurement from $\mathbf{y}$. These measurement matrices are referred as *dense*. In some practical cases, the measurements may depend only on some of the elements from $\mathbf{X}$, and the elements that do not contribute to the formation of the measurements are designated by zero values in the measurement matrix, $\mathbf{A}$ [11]–[14]. These measurement matrices are, therefore, sparse and may also be well structured. Typically, two approaches are used for their processing, of which one is based on the multiplication of the original measurement matrix, $\mathbf{A}$, by a random $N \times N$ matrix, $\mathbf{B}$, to yield a dense measurement matrix, $\mathbf{AB}$. The other approach employs the sparse and structured measurement matrices in their original form for the reconstruction.

As a simple example of the latter approach, consider the structured, partial, nonoverlapping, short-time Fourier transform measurement matrix [14]. The full measurement matrix, $\mathbf{A}$, consists of blocks of the partial DFT matrices, so the dimensionality for the analysis may be easily reduced to the examination of the constituting blocks. It is important to notice that each block must satisfy the general CS conditions, such as those presented in this article.

Detection of unknown coefficient positions
In the CS setup, the positions of the nonzero elements in $\mathbf{X}$ are typically not known. A natural approach to solve the CS reconstruction problem would be to adopt a two-step strategy as follows:
1) detect the positions of the nonzero elements
2) apply a reconstruction algorithm with the known positions of the nonzero elements.

An intuition for the estimate of the positions of the nonzero elements in the first step comes from the linear nature of the measurements, $y(m)$, obtained as linear combinations of the sparsity domain elements, $X(k)$, and the corresponding rows of the measurement matrix, $\mathbf{A}$.

The Direct Search Solution

The simplest and most intuitive way to extend the solution to compressive sensing (CS) that has unknown positions in X, as in (3), to the general case with K unknown nonzero positions, $k_1, k_2, \ldots, k_K$, would be to consider all possible combinations of the K nonzero positions (out of N) and solve the system in (3) for each combination that is dictated by the sequence $k_1, k_2, \ldots, k_K$. After solving the system in the least-squares sense, the solution should be checked against the given measurement in (3). A zero error indicates a successful solution of the reconstruction problem. The solution is, therefore, unique if only one combination of the coefficients $k_1, k_2, \ldots, k_K$ produces zero error.

A practical problem with this kind of reconstruction is that it requires the evaluation of $\binom{N}{K}$ combinations, a nondeterministic polynomial-time hard problem with an unaffordable computational burden. For example, for $N = 1{,}024$ and $K = 5$, we would require 10^{13} combinations, whereby the corresponding overdetermined system (3) needs to be solved for each combination.

The Restricted Isometry Property Calculation

A very popular approach to check, in theory, for the uniqueness of the solution is based on the restricted isometry property (RIP) of the measurement matrix. This can be formalized by stating that a K-sparse solution is unique if measurement matrix A satisfies the RIP condition with a constant $0 \leq \delta_{2K} < 1$ for a $2K$-sparse signal, while to obtain the RIP constant, a check of a $2K$-sparse signal requires $\binom{N}{2K}$ combinations of the submatrices, $\mathbf{A}_{2K}$, and their eigenvalue decomposition. Therefore, this is an even more complex problem than the direct search solution,

which makes it prohibitive for practical use. For example, for a vector with $N = 1{,}024$ entries out of which $K = 5$ nonzero coefficients are considered, the RIP check requires 10^{23} combinations and matrix calculations. To put this into perspective, for a given measurement matrix, the RIP is checked only once; however, if the same matrix is used in not more than 10^{10} experiments, the direct combinatorial search is still more efficient, although it is equally computationally prohibitive and often not feasible in practice.

The Complexity of the Coherence Index Calculation

The coherence index check requires only $\binom{N}{2}$ combinations, which is both computationally feasible and sparsity invariant. This makes it a preferred choice in the optimization of measurement matrices. Various other efforts to avoid RIP calculation in the uniqueness check may be found in the literature, among which the most notable are those referred to as the RIP-less approaches [S1]–[S4].

References

[S1] H. Arguello and G. R. Arce, "Colored coded aperture design by concentration of measure in com- pressive spectral imaging," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1896–1908, Apr. 2014. doi: 10.1109/TIP.2014.2310125.

[S2] E. J. Candes and Y. Plan, "A probabilistic and RIPless theory of compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7235–7254, Nov. 2011. doi: 10.1109/TIT.2011.2161794.

[S3] H. Calderón, J. F. Silva, J. M. Ortiz, and A. Egaña, "Reconstruction of channelized geological facies based on RIPless compressed sensing," *Comput. Geosci.*, vol. 77, pp. 54–65, Apr. 2015. doi: 10.1016/j.cageo.2015.01.006.

[S4] L. Stanković and M. Daković, "On the uniqueness of the sparse signals reconstruction based on the missing samples variation analysis," *Math. Probl. Eng.*, vol. 2015, 2015. doi: 10.1155/2015/629759.

Remark 2

The linearity of the CS paradigm in (2) admits a back-projection of the measurements, $\mathbf{y}$, to the measurement matrix, $\mathbf{A}$, defined by

$$\mathbf{X}_0 = \mathbf{A}^H \mathbf{y} = \mathbf{A}^H \mathbf{A} \mathbf{X}, \qquad (6)$$

to estimate the positions of the nonzero elements in $\mathbf{X}$. In an ideal case, the matrix $\mathbf{A}^H \mathbf{A}$ should ensure that the initial estimate, $\mathbf{X}_0$, contains exactly K elements at positions $\{k_1, k_2, \ldots, k_K\}$, for which the magnitudes are greater than the largest one at the remaining positions. Then, by taking the positions of the highest-magnitude elements in $\mathbf{X}_0$ as the set $\{k_1, k_2, \ldots, k_K\}$ in (3), the algorithm

for the known nonzero element positions, from the previous section, can be applied to reconstruct the signal.

Remark 3

Note that if $\mathbf{A}^H \mathbf{A}$ were an identity matrix, the initial estimate, $\mathbf{X}_0$, would correspond to the exact solution, $\mathbf{X}$. However, with a reduced number of measurements, $M < N$, this cannot be achieved (due to the Welch lower bound). A pragmatic requirement for the existence of the CS solution would, therefore, be that the off-diagonal elements of $\mathbf{A}^H \mathbf{A}$ are as small as possible compared to the unit diagonal elements.

The condition that all of the K elements in the initial estimate, $\mathbf{X}_0$, that are

located at the nonzero positions in the original sparse vector, $\mathbf{X}$, are larger than any other element in the initial estimate, $\mathbf{X}_0$, can be relaxed through an iterative procedure. To be able to find the position, k_1, of the largest nonzero element in $\mathbf{X}_0$, its value must be larger than the values $X_0(k)$ at the original zero-valued positions. After the position of the largest element is found and its value estimated, this component can be reconstructed and subtracted from the measurements, $\mathbf{y}$, and the procedure is continued with the remaining $(K-1)$-sparse elements in an iterative manner. The stopping criterion becomes that $\mathbf{A}_{MK} \mathbf{X}_K = \mathbf{y}$ should hold for the estimated nonzero positions

$\{k_1, k_2, \ldots, k_K\}$ and elements $X(k)$, as outlined in Algorithm 1.

Practical guidelines

In real-world scenarios, when we want to apply CS methods to the reconstruction of an N-dimensional signal from its reduced M-dimensional measurements, signal sparsity must be assumed, since it is a prerequisite for any CS-based reconstruction. However, for real-world signals, an exact sparsity, K, is rarely

known in advance, meaning that the M measurements for the signal reconstruction problem should be empirically assumed (for example, the signal bandwidth in classical signal sampling). After each step during iterative reconstruction, as in Algorithm 1, we may easily test if the stopping criterion, $\mathbf{A}_{MK}\mathbf{X}_K = \mathbf{y}$, is satisfied, as it indicates that the true sparsity degree has been reached and detected. We should subsequently check for the uniqueness of the solution using the appropriate criteria to ensure that there are no other possible solutions with the same sparsity. If the stopping criterion is not reached after the sparsity degree, K, reaches the maximal expected sparsity, the number of measurements, M, or the structure of the measurement matrix, $\mathbf{A}$, does not satisfy the reconstruction conditions, the matrix must be redesigned, or the number of measurements increased.

The unique reconstruction condition

The key criterion for faithful signal reconstruction from a reduced set of measurements is the uniqueness of the result. In CS methodology, the uniqueness is commonly defined through the coherence index, as stated below.

Proposition

The reconstruction of a K-sparse signal, $\mathbf{X}$, is unique if the coherence index, μ, of the measurement matrix, $\mathbf{A}$, satisfies [2]

$$K < \frac{1}{2}\left(1 + \frac{1}{\mu}\right), \qquad (7)$$

where the coherence index is defined by

$$\mu = \max_{\substack{k,l \\ k \neq l}} \left| \frac{\sum\limits_{m=0}^{M-1} a_m(k)\, a_m^*(l)}{\sqrt{\sum\limits_{m=0}^{M-1} |a_m(k)|^2 \sum\limits_{m=0}^{M-1} |a_m(l)|^2}} \right|. \qquad (8)$$

For a normalized measurement matrix, $\mathbf{A}$, the coherence index, μ, is equal to the maximum absolute off-diagonal element of $\mathbf{A}^H \mathbf{A}$.

The proposition is usually proved based on the Gershgorin disk theorem [10], a topic that is not covered in the engineering curricula; this obstacle prevents the more widespread engagement of engineers in the CS field and wider success of the eventual applications. To this end, we proceed to derive the reconstruction condition in a self-contained and intuitive way that does not require advanced mathematics.

As an example, consider a 7×14 measurement matrix, $\mathbf{A}$; for clarity, we
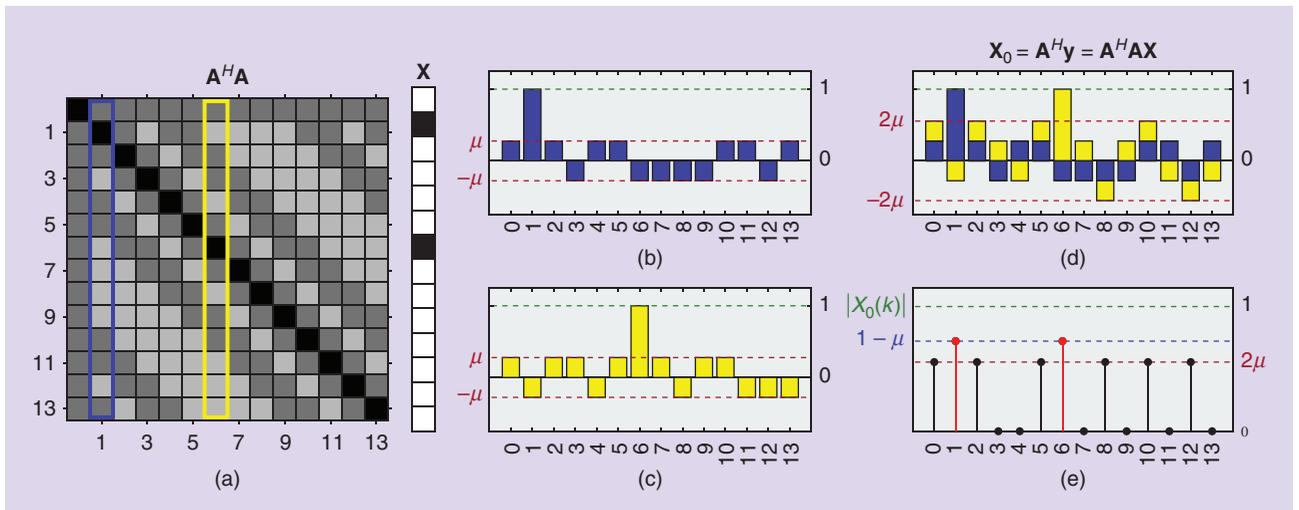

FIGURE 2. An illustration of the coherence index relation for $K = 2$. (a) The matrix $\mathbf{A}^H\mathbf{A}$ for the ETF matrix $\mathbf{A}$ from Figure 1 and the original elements in $\mathbf{X}$ with nonzero values at $k_1 = 1$ and $k_2 = 6$. The values in matrix $\mathbf{A}^H\mathbf{A}$ are indicated by black equals 1, white equals 0, dark gray equals $+\mu$, and light gray equals $-\mu$. The components in the initial estimate $\mathbf{X}_0 = \mathbf{A}^H\mathbf{y}$, calculated using the measurements $\mathbf{y} = \mathbf{AX}$ that result from (b) the nonzero element at $k_1 = 1$ in the blue bar plot and (c) $k_2 = 6$. in the yellow bar plot. The visual demonstration that the condition for a correct detection of the original nonzero element position from the initial estimate is $1 - \mu > 2\mu$, with (d) the stacked bar plot for both components and (e) the resulting $|X_0(k)|$.

employ the so-called equiangular tight frame (ETF) matrix, for which the absolute value of the off-diagonal elements of $\mathbf{A}^H\mathbf{A}$ is constant and equal to the coherence index, μ. Matrix $\mathbf{A}^H\mathbf{A}$ is visualized in Figure 2(a). Its diagonal elements are, by definition, equal to one, while its off-diagonal elements that have values of $\pm\mu$ can be treated as the disturbances in the detection of the nonzero element positions in $\mathbf{X}$. Observe that the off-diagonal column elements of this matrix represent the normalized contribution of the corresponding nonzero element in the sparse vector, $\mathbf{X}$, to the cumulative disturbance.

Consider the case with only one nonzero element in $\mathbf{X}$, say at position $k_1 = 1$, with its initial estimate, $\mathbf{X}_0$, designated by the dark blue bar plot in Figure 2(b). For only one nonzero element in $\mathbf{X}$, the condition for the correct detection of its position in the estimate $\mathbf{X}_0$ would be that the maximum possible disturbance value, μ, is smaller than the value on the diagonal, which is, in this case, $\mu < 1$.

Assume now that the signal sparsity index is $K = 2$, with nonzero elements in $\mathbf{X}$ at $X(1) = 1$ and $X(6) = 1$, as indicated in Figure 2(a) by the respective yellow and blue columns in matrix $\mathbf{A}^H\mathbf{A}$. The set of $M = 7$ measurements, $\mathbf{y} = \mathbf{AX}$, is used to calculate the initial estimate in (6) as $\mathbf{X}_0 = \mathbf{A}^H\mathbf{y} = \mathbf{A}^H\mathbf{AX}$, as shown in the stacked bar plot in Figure 2(d). More specifically, the multiplication of $\mathbf{A}^H\mathbf{A}$ by $\mathbf{X}$ results in two components in the initial estimate:

■ The component whose values are shown in dark blue bars represents the contribution of $X(1)$ to the initial estimate $\mathbf{X}_0$.
■ The component designated by yellow bars represents the corresponding contribution of $X(6)$.

Remark 4

Compared to the case where the sparsity is $K = 1$, we can observe two differences when the sparsity degree is $K = 2$:

■ The disturbances arising from each nonzero element in $\mathbf{X}$ combine so that the maximum possible disturbance is increased to 2μ.
■ The unit value of the original nonzero element in $\mathbf{X}$ is also affected by the disturbing values of the other nonzero element, with the maximum possible amplitude reduction of this nonzero element of $(1 - \mu)$, as shown in the stacked bar plot in Figure 2(d).

Remark 5

From the absolute value of the initial estimate in Figure 2(e), that is, |"blue bars" + "yellow bars"|, we can conclude that for $K = 2$, the correct nonzero element position in $\mathbf{X}$ will always be detected if the original unit amplitude, reduced by the maximum possible disturbance, μ, is greater than the maximum possible disturbance value, 2μ, at the original zero-valued positions in $\mathbf{X}$. In other words, for $K = 2$,

the reconstruction condition is given by $1 - \mu > 2\mu$. If this condition is met, the position of the nonzero element will always be correctly detected, which is precisely the aim of uniqueness analysis. Note that, for rigor, we assumed the worst-case scenario for the largest position detection of $X(1) = X(6) = 1$, whereby the disturbance from the other nonzero element was the strongest possible. If $|X(1)| > |X(6)|$, it would relax the condition for detection of the $|X(1)|$ position, and vice versa.

More specifically, the coherence index for the considered ETF matrix is $\mu = 0.2774$. Therefore, in the worst-case scenario, the initial estimated values at the nonzero positions $k_1 = 1$ and $k_2 = 6$ would be $1 - \mu = 1 - 0.2774 = 0.7226$, which is greater than the largest possible value, $2\mu = 2 \times 0.2774 = 0.5547$, of the initial estimate at the positions where the original vector $\mathbf{X}$ is zero-valued, $k \notin \{1, 6\}$. Therefore, the coherence index condition $K < 0.5(1 + 1/\mu) = 2.3$ is satisfied for $K = 2$. Observe, also, that matrix $\mathbf{A}$, defined previously, cannot be used with $K \geq 3$, since the maximum possible disturbance of 3μ would be larger than a maximally reduced unit value at the nonzero element position in the initial estimate. The amount of this maximum reduction would be 2μ, resulting in the initial estimate value of $(1 - 2\mu)$ and a misdetection of the nonzero element position in $\mathbf{X}$.

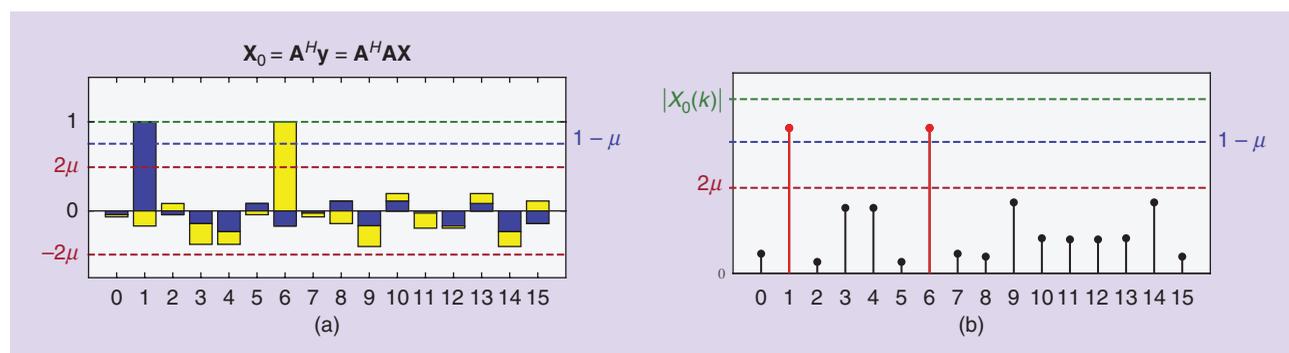As an example of the tightness of the coherence index condition, Figure 3

FIGURE 3. (a) The initial estimate, $\mathbf{X}_0$, for a 12×16 partial DFT measurement matrix, where $\mu = 0.2455$, and the sparsity degree of $K = 2$ in $\mathbf{X}$, with $X(1) = X(6) = 1$ [the real part of $X_0(k)$ is presented in the stacked bar plot]. (b) The positions of nonzero elements in $\mathbf{X}$ are detected at $k \in \{1, 6\}$ if their amplitude, when reduced by the maximum possible disturbance from the other component, μ, is above the maximum possible disturbance, 2μ, of both components, at a position of an original zero-valued element in $\mathbf{X}$, $k \notin \{1, 6\}$. Observe that there is a larger misdetection margin compared to the ETF case, indicated by the corresponding $(1 - \mu)$ and 2μ levels, that is a consequence of different distributions of disturbances (off-diagonal elements of matrix $\mathbf{A}^H\mathbf{A}$).
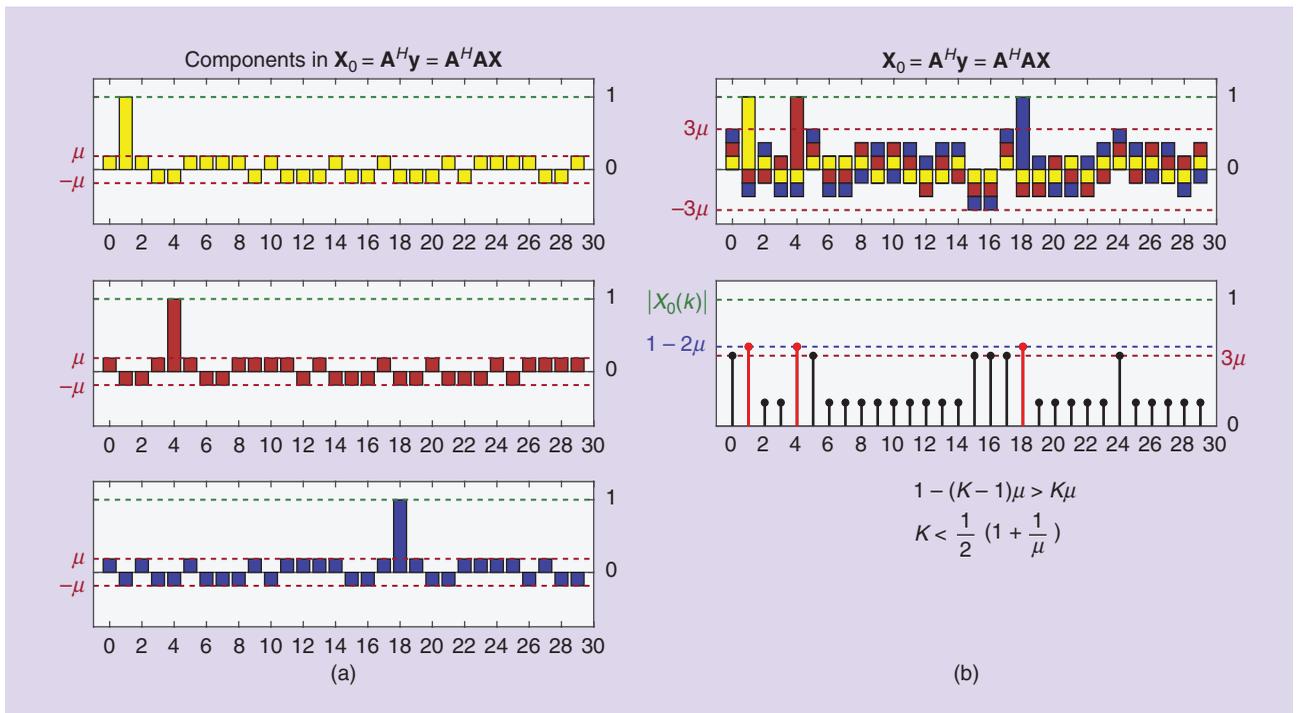
FIGURE 4. The illustration of the coherence index relation for $K = 3$. (a) The components of the initial estimate, $\mathbf{X}_0 = \mathbf{A}^H \mathbf{y} = \mathbf{A}^H \mathbf{A} \mathbf{X}$, that result from each of the three unity-valued nonzero elements in $\mathbf{X}$ at $k \in \{1, 4, 18\}$. (b) The components are combined into the initial estimate, $\mathbf{X}_0$. The positions of the nonzero elements in $\mathbf{X}$ are detected at $k \in \{1, 4, 18\}$ using the absolute value of the initial estimate, $|X_0(k)|$, if their amplitudes, reduced for a maximum possible disturbance, 2μ, from the other two components are above the maximum possible disturbance, 3μ, of all three components at the position of an original zero-valued element in $\mathbf{X}$, $k \notin \{1, 4, 18\}$. The condition for the correct detection of the nonzero element positions with $K = 3$ is, therefore, $1 - 2\mu > 3\mu$; that is, $1 - (K - 1)\mu > K\mu$, which generalizes to any K.

presents the CS recovery based on a 12×16 partial DFT measurement matrix, $\mathbf{A}$, with $\mu = 0.2455$ for the case of $K = 2$. Observe that, unlike the ETF case, the disturbance terms are not equal. However, all of the conclusions regarding the worst-case scenario remain valid, and the nonzero element positions in $X(k)$ will be correctly detected based on the presented initial estimate, $X_0(k)$, if $1 - \mu > 2\mu$. To provide further intuition, the illustration in Figure 2 is repeated for a 15×30 measurement matrix, $\mathbf{A}$, and $K = 3$, with the results shown in Figure 4. Following the same reasoning as in the previous case, we can conclude from the absolute value of the initial estimate that the detection condition now becomes $1 - 2\mu > 3\mu$. The coherence index for this matrix is, therefore, $\mu = 0.1857$, with the corresponding coherence index condition $K < 0.5(1 + 1/\mu) = 3.2$. Notice that this matrix cannot be used for $K = 4$, since $1 - 3\mu > 4\mu$ would not hold, meaning that a disturbance would be incorrectly detected as a desired nonzero component.

Remark 6

Following the simple and inductive approach, it becomes immediately obvious that, for a general case of a K-sparse $\mathbf{X}$, the position of the largest element position in $\mathbf{X}$ will be correctly detected in the initial estimate, $\mathbf{X}_0$, if

$$1 - (K - 1)\mu > K\mu.$$

The bound directly yields the coherence index condition in (7), with the derivation obtained in a natural and practically relevant way.

Remark 7

After the position of the first nonzero component in a sparse $\mathbf{X}$ is successfully detected, reconstructed, and removed, the same procedure and relations can be iteratively applied to the remaining "deflated" signal, which now exhibits a reduced $(K - 1)$-sparsity level, thus guaranteeing a unique solution. A simple analytic derivation to support the intuition behind the proposed proof of the condition for unique

reconstruction is provided in "The Proposed Derivation of the Coherence Index Relation in (7)," and as illustrated in Figures 2–4. The case with small measurement matrices of the ETF type, which have been considered so far, fully supports the proposed derivation of the coherence index and its appropriateness in theory and practice, as it produces a tight bound on the existence and uniqueness of the reconstruction. The corresponding performance for problems of large dimensions is illustrated in "The Coherence Index: From Theory to Practice."

In summary, the coherence index has been derived for several independent worst-case scenarios to provide a rigorous and easily interpretable uniqueness bound. In practical applications, the uniqueness condition is typically further relaxed. To this end, in our approach, we considered various cases.

■ The *amplitudes of all of the nonzero components in* $\mathbf{X}$ *are equal*: If this is relaxed to the general case of different values of the nonzero elements in

X, we can expect a successful, more relaxed, and unique reconstruction, even if the coherence index condition may be violated.

- *Only one element in the initial estimate, X_0, was compared to the maximum possible disturbance*: Having any one of the original nonzero components in the initial estimate at a position of a nonzero element in **X** above the maximum disturbance would guarantee a successful reconstruction and further relax the reconstruction condition.

- *All of the disturbances at the nonzero- and zero-valued element positions are assumed to be in phase*: That is, they sum together with the maximum possible magnitudes. This is a very low-probability event that, again, relaxes the practical sparsity limit for the reconstruction.

- *The distribution of the off-diagonal elements of $A^H A$ plays an important role*: When there are several nonzero elements in **X** they combine to an approximately Gaussian distributed variable, and the resulting disturbance may obey different distributions and yield varying results.

For large sparsities, K, the misdetection probability (shown in Figure S1) can be improved (lowered) by increasing the upper limit for the iterations in Algorithm 1 by a few percentage points, with respect to the expected sparsity, K. After the iterations are completed, the expected sparsity, K, is used in the final reconstruction. This kind of calculation solves the problem that results from the fact that the iterative reconstruction in Algorithm 1 cannot produce the exact result if it misses one of the nonzero coefficient positions.

A real-world example

As an example, we considered the test image "pout" from MATLAB. When compressed in the 2D discrete cosine transform (2D-DCT) domain, this image may be considered as sparse. In our experiment, it was initially compressed to retain the eight largest 2D-DCT values in each 8×8 block of pixels (a 12.5% compression ratio). Next, 40% of the pixels were corrupted with "salt and pepper" noise and marked as missing/unavailable, and they are designated in white in Figure 5(a). The image was reconstructed from this incomplete version using Algorithm 1, with the available pixels serving as the measurements, **y**, and the partial 2D-DCT matrix as the measurement matrix, **A**. The coherence index was calculated as the maximum absolute off-diagonal element in $A^H A$, as in Figure 2. Figure 5(c)

The Proposed Derivation of the Coherence Index Relation in (7)

The formal definition of the coherence index is

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

where it is assumed that the measurement matrix is normalized; that is, $\sum_{m=0}^{M-1} |a_m(k)|^2 = 1$ for each k.

A measurement (or signal sample) of a K-sparse signal can be written, by definition, as

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

Its initial estimate in (6), for one coefficient, $X_0(k)$, is

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

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

Without a loss of generality, assume that the element $X(k_1)$ at position k_1 is real-valued with the unit amplitude, $X(k_1) = 1$, while the remaining elements are not greater than the unity, $|X(k_i)| \leq 1$, $i = 2, \ldots, K$. The unit value at k_1 will be disturbed by the influence of other nonzero coefficients; that is,

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

In the worst-case scenario for the detection of the coefficient at position k_1, all $X(k_i) \mu(k, k_i)$ should assume their lowest possible value of $-\mu$. Then,

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

and the minimum possible value of $|X_0(k_1)|$, therefore, becomes

$$\min |X_0(k_1)| = 1 - (K-1)\mu.$$

For correct detection, the amplitude of $X_0(k_1)$ should be greater than the maximum possible disturbance at positions where the original coefficients are zero-valued, $k \neq k_i$, which is equivalent to

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

The condition that the coefficient position, k_1, must be correctly detected becomes

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

From the last two relations, it immediately follows that

$$1 - (K-1)\mu > K\mu,$$

which, as desired, produces the coherence index relation given in (7).

To establish a link with common compressive sensing practice, we statistically compare the reconstruction performance based on the following measurement matrices: the equiangular tight frame (ETF), Gaussian matrix (real- and complex-valued with identical and independently distributed real and imaginary parts), and partial discrete Fourier transform (DFT) measurement matrix, with $N = 258$ and $M = 129$.

For the Gaussian and partial DFT matrices, we chose the best measurement matrix (with the smallest μ) from 1,000 random realizations. The corresponding values of μ were 0.3527, 0.2737, 0.0624, and 0.1484 for the Gaussian (real- and complex-valued) matrix, ETF, and partial DFT matrix, respectively. The lowest value for the coherence index, $\mu = \sqrt{(N - M)/(M(N - 1))} = 0.0624$, was obtained for the ETF, as expected, from the Welch bound theory. The theoretic limits for the sparsity degree, K, that guarantee the reconstruction with a probability of one are $K < 0.5(1 + 1/\mu) = 8.5$, $K < 3.9$, $K < 2.3$, and $K < 1.9$ for the ETF, partial DFT, and Gaussian (complex- and real-valued) measurement matrices, respectively.

In the experiment, all four matrices were used to reconstruct signals for a range of sparsity degrees, $K = 1, 2, 3, \ldots, 64$. For each sparsity level, K, the problem was solved 100,000 times with random positions of the nonzero coefficients. For each K, the solution was checked against the known positions and values of the nonzero elements in $X(k)$, and the number of mis-detections was recorded. The number of misdetections for each K was then divided by the total number of realizations.

The results for the misdetection probability are shown in the Figure S1. For a sufficient practical probability value of the misdetection (for example, lower than 10^{-4}), the partial DFT performed significantly better than the ETF. Figure S1 shows that it produced a reconstruction without any error, with a probability of 0.9999, for $K < 21$, while for the ETF, the result holds for $K < 15$.

This can be explained by the probability distribution of the off-diagonal elements of $\mathbf{A}^H\mathbf{A}$ being the worst possible in the ETF, with $p(\xi) = 0.5(\delta(\xi - \mu) + \delta(\xi + \mu))$, and the variance of $\sigma^2 = \mu^2 = (N - M)/(M(N - 1))$. For the partial DFT matrix, the off-diagonal elements of $\mathbf{A}^H\mathbf{A}$ were approximately Gaussian distributed, with the variance $\sigma^2 = (N - M)/(M(N - 1))$ [S5], while in the Gaussian measurement matrix, the off-diagonal elements were also Gaussian distributed, with $\sigma^2 = 1/M$.

The probabilistic approach can also be used to derive a relation between K, N, and M for a successful reconstruction of $\mathbf{X}$. with a given probability, as thoroughly shown in [S5]. We present a brief practical analysis. The real and imaginary parts of the off-diagonal elements of $\mathbf{A}^H\mathbf{A}$ (the disturbances at the original zero-valued positions in $\mathbf{X}$) combine within the initial estimate for each nonzero element in $\mathbf{X}$. For a large sparsity degree, $1 \ll K \ll N$, the resulting distribution of the disturbance is approximately Gaussian, $\propto \mathcal{N}(0, K\sigma^2/2)$, where σ^2 is the variance of a single complex-valued, off-diagonal element of $\mathbf{A}^H\mathbf{A}$. A successful reconstruction may be expected if the amplitude of the disturbance (Rayleigh distributed) is, with high probability, below the nonzero unit-amplitude element in the initial estimate. For example, if $K\sigma^2$ is denoted by $1/\kappa$, the amplitude in at least one position of the zero-valued elements in $\mathbf{X}$ will be above one, with a probability of $(N - K)e^{-\kappa} \sim Ne^{-\kappa}$. For a partial DFT matrix, assuming $\kappa = 16$ when $Ne^{-\kappa} \sim 10^{-5}$, the reconstruction is achieved with a high probability for

$$K < \frac{1}{16} \frac{M(N - 1)}{N - M}.$$

For $N = 258$ and $M = 129$, it follows that $K < M/8 = 16.1$, as suggested for the practical sparsity limit in [9]. This corresponds to the previous statistical results, with a probability of incorrect reconstruction below 10^{-5}.

FIGURE S1. The misdetection probability of the original nonzero element positions, given in (a) linear and (b) logarithmic scales. The measurement matrix, **A**, took the forms of ETF (red); random, partial DFT (black); random, real-valued Gaussian (green); and complex-valued Gaussian measurement (blue) matrices.

Reference

[S5] L. Stanković, S. Stanković, and M. G. Amin, "Missing samples analysis in signals for applications to L-estimation and compressive sensing," *Signal Process.*, vol. 94, pp. 401–408, Jan. 2014. doi: 10.1016/j.sigpro.2013.07.002.
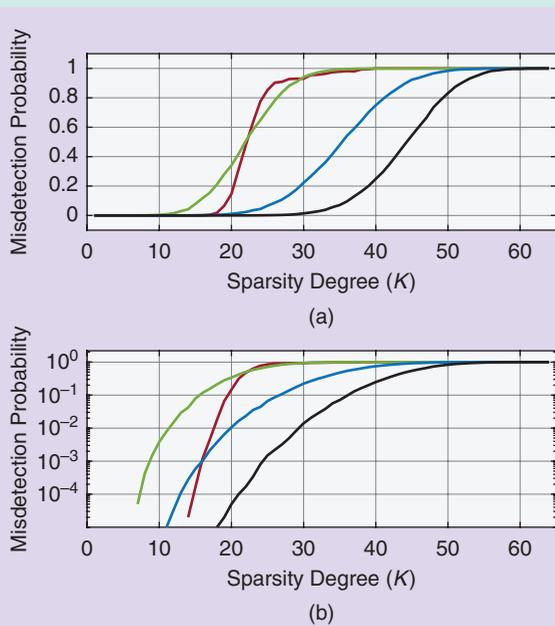
and (d) illustrates the reconstruction procedure and shows the measurement matrix, $\mathbf{A}$, values of the 2D-DCT for one 8×8 pixel block arranged into vector form, $\mathbf{X}$, and the available pixels for the same block, $\mathbf{y}$. The achieved mean square error of the reconstructed image, compared to the original before compression, was –34 dB, while the achieved peak signal-to-noise ratio with respect to the original (with all of the pixels available and before compression) was 86 dB.

If we were to check the RIP condition for this example, we would have to examine all of the submatrices that were obtained after choosing $2K = 16$ columns from matrix $\mathbf{A}$. The number of submatrices would be

$$\binom{64}{16} = 4.9 \times 10^{14},$$

and for each submatrix, eigen decomposition must be performed to obtain the corresponding eigenvalues. The largest eigenvalue should be selected for the RIP constant for $2K = 16$, denoted by δ_{16}, and checked to see if it is below the theoretical limits. This exemplifies that, even for low values of $N = 64$ and $K = 8$, the RIP calculation is not feasible. To obtain the coherence index, we have

$$\binom{64}{2} = 2,106$$

combinations of the columns. The same reconstruction framework can be applied to many other practical scenarios, including those presented in [7], [13], and [15].

Conclusions

The coherence index condition for unique sparse signal reconstruction, which is a prerequisite for the successful CS paradigm, has been derived using simple signal processing tools and through an intuitive example. Our perspective has demonstrated that this index provides a tight uniqueness bound for relatively small dimensions of CS problems, followed by a clear interpretation of its conservative nature for large-scale issues. This has been achieved for a range of high probabilities for obtaining the correct result and avoiding misdetection. It
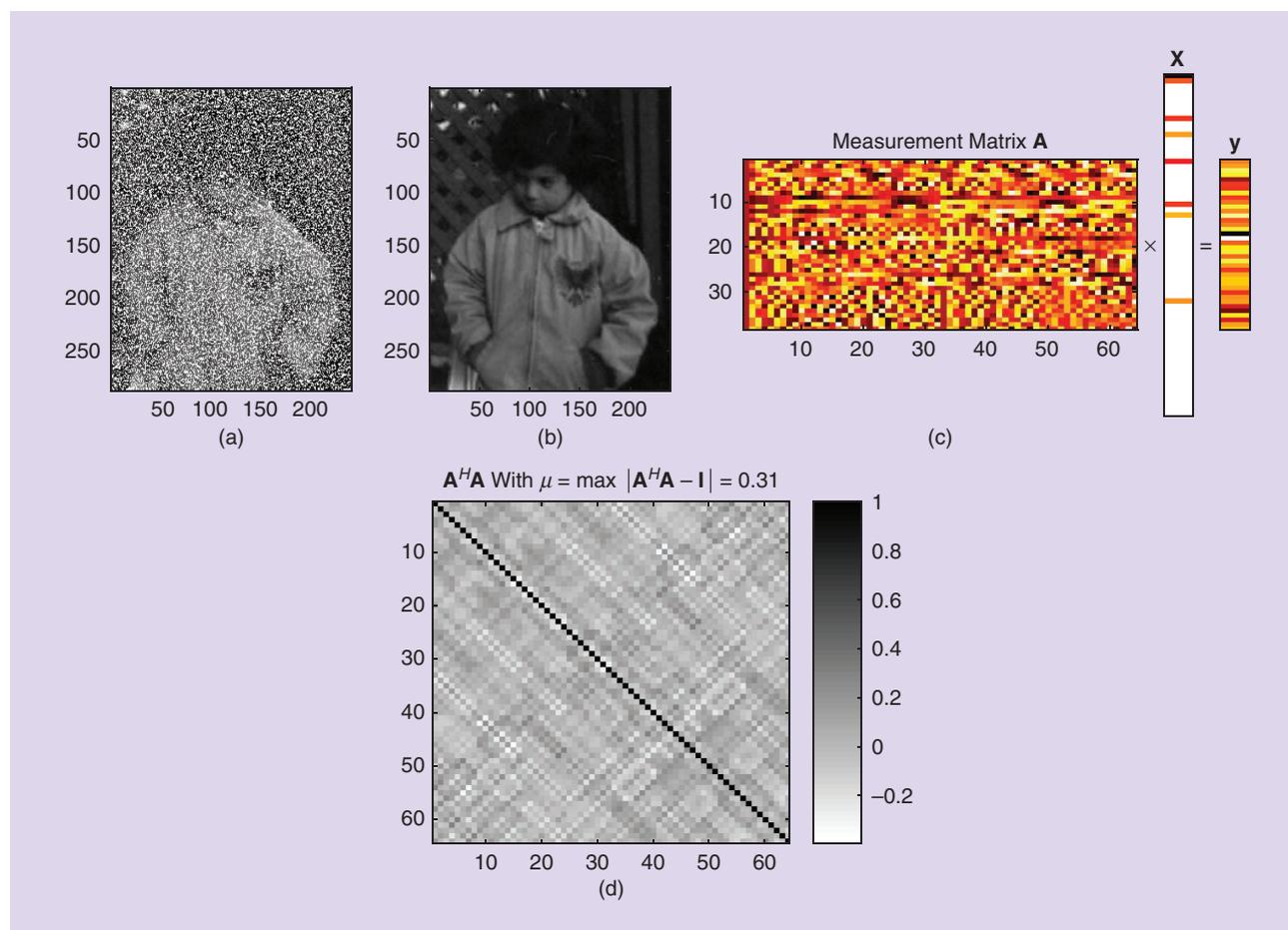

FIGURE 5. The reconstruction of a compressed image (sparse in the 2D-DCT domain) with 40% of the missing/unavailable/corrupted pixels. (a) The original image, with the missing pixels designated in white. (b) The reconstructed image obtained using Algorithm 1. (c) The measurement matrix, $\mathbf{A}$; the sparse coefficients of the 2D-DCT domain, $\mathbf{X}$; and the observable measurements, $\mathbf{y}$, for one typical 8×8 block of the image. (d) The matrix $\mathbf{A}^H\mathbf{A}$, whose maximum absolute off-diagonal element is equal to $\mu = 0.31$, which ensures a unique recovery for sparsity degrees $K \leq 2$ (see Figure 2). However, with the probabilistic effects (explained in "The Coherence Index: From Theory to Practice"), the recovery was of high quality, even with $K = 8$, as used in this example. The achieved mean square error was −34 dB, while the peak signal-to-noise ratio was 86 dB (with all of the pixels available and before compression). This demonstrates the rather conservative nature of the coherent index that is bound in practical applications and supports the analysis in "The Coherence Index: From Theory to Practice."

was also shown that general measurement matrices, such as the frequently used partial DFT matrix, are likely to outperform results derived from the considered ETF-based measurement matrix as long as the analysis is restricted to pragmatically high probabilities of the correct solution, resulting in a practical relaxation of the theoretical probability of one for avoiding misdetection.

Authors

Ljubiša Stanković (ljubisa@ac.me) is a professor with the University of Montenegro, Podgorica. He is the recipient of the 2017 European Association for Signal Processing Best Journal Paper Award. He is a Fellow of the IEEE and a member of the Montenegrin Academy of Sciences and Arts and the European Academy of Sciences.

Danilo P. Mandic (d.mandic@ imperial.ac.uk) is a professor of signal processing at Imperial College London, United Kingdom. He is the 2019 recipient of the Dennis Gabor Award of the International Neural Networks Society, the 2018 IEEE Signal Processing Best Paper Award, and the President's Award for Excellence in Postgraduate Supervision at Imperial College in 2014. He is a Fellow of the IEEE.

Miloš Daković (milos@ac.me) received his B.Sc., M.S., and Ph.D. degrees from the University of Montenegro, Podgorica, in 1996, 2001, and 2005, respectively, and is a professor at the same university. His research interests are in signal processing and computer science. He is a Member of the IEEE.

Ilya Kisil (i.kisil15@imperial.ac.uk) received his M.Sc. degree in intelligence systems in robotics from ITMO University, Saint Petersburg, Russia, and his M.Res. degree in advanced computing from Imperial College London, United Kingdom, where he is a Ph.D. candidate. His research interests include tensor decompositions, big data, efficient software for large-scale problems, and graph signal processing.

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. doi: 10.1109/MSP.2007.914731.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006. doi: 10.1109/TIT.2006.871582.

[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. doi: 10.1007/s00034-018-0909-2.

[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. doi: 10.1109/TIT.2005.860430.

[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. doi: 10.1109/TIT.2007.909108.

[6] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *C. R. Math.*, vol. 346, no. 9–10, pp. 589–592, May 2008. doi: 10.1016/j.crma.2008.03.014.

[7] L. Stanković, "On the ISAR image analysis and recovery with unavailable or heavily corrupted data," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 3, pp. 2093–2106, July 2015. doi: 10.1109/TAES.2015.140413.

[8] L. Stanković and M. Brajović, "Analysis of the reconstruction of sparse signals in the DCT domain applied to audio signals," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 7, pp. 1216–1231, July 2018. doi: 10.1109/TASLP.2018.2819819.

[9] 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. doi: 10.1109/TIT.2005.862083.

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

[11] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Appl. Opt.*, vol. 47, no. 10, pp. B44–B51, Apr. 2008. doi: 10.1364/AO.47.000B44.

[12] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "Compressive coded aperture spectral imaging: An introduction," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 105–115, Jan. 2014. doi: 10.1109/MSP.2013.2278763.

[13] L. Wang, Z. Xiong, D. Gao, G. Shi, and F. Wu, "Dual-camera design for coded aperture snapshot spectral imaging," *Appl. Opt.*, vol. 54, no. 4, pp. 848–858, Feb. 2015. doi: 10.1364/AO.54.000848.

[14] L. Stanković, I. Orović, S. Stanković, and M. Amin, "Compressive sensing based separation of nonstationary and stationary signals overlapping in time-frequency," *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4562–4572, Sept. 2013. doi: 10.1109/TSP.2013.2271752.

[15] I. Stanković, and W. Dai, "Reconstruction of global ozone density data using a gradient-descent algorithm," in *Proc. 2016 Int. Symp. ELMAR*, Zadar, Croatia. doi: 10.1109/ELMAR.2016.7731760.

SP

READER'S CHOICE

less runtime without compromising the clustering quality.

2016

Deduplication on Encrypted Big Data in Cloud

Yan, Z.; Ding, W.; Yu, X.; Zhu, H.; Deng, R.H.

A scheme based on data ownership challenge and proxy re-encryption is proposed to manage encrypted data storage with deduplication. It supports flexible data update and sharing with deduplication, even when the data holder is offline, and integrates cloud data deduplication with access control. Extensive performance analysis and tests show the superior efficiency and effectiveness of the scheme for potential practical deployment, especially for big data deduplication in cloud storage.

2016

Petuum: A New Platform for Distributed Machine Learning on Big Data

Xing, E. P.; Ho, Q.; Dai, W.; Kyu Kim, J.; Wei, J.; Lee, S.; Zheng, X.; Xie, P.; Kumar, A.; Yu, Y.

This paper proposes a general-purpose framework, Petuum, that systematically addresses data- and model-parallel challenges in large-scale machine learning. It observes that many machine-learning programs are fundamentally optimization centric and admit error-tolerant, iterative-convergent algorithmic solutions. It presents unique opportunities for integrative system design, such as bounded-error network synchronization and dynamic scheduling based on machine-learning program structure.

2015

SP