RECONSTRUCION OF SPARSE AND NONSPARSE SIGNALS FROM A REDUCED SET OF SAMPLES

Ljubiša Stanković, Isidora Stanković

Abstract: Signals sparse in a transformation domain can be recovered from a reduced set of randomly positioned samples by using compressive sensing algorithms. Simple reconstruction algorithms are presented in the first part of the paper. The missing samples manifest themselves as a noise in this reconstruction. Once the reconstruction conditions for a sparse signal are met and the reconstruction is achieved, the noise due to missing samples does not influence the results in a direct way. It influences the possibility to recover a signal only. Additive input noise will remain in the resulting reconstructed signal. The accuracy of the recovery results is related to the additive input noise. Simple derivation of this relation is presented. If a reconstruction algorithm for a sparse signal is used in the reconstruction of a nonsparse signal then the noise due to missing samples will remain and behave as an additive input noise. An exact relation for the mean square error of this error is derived for the partial DFT matrix case in this paper and presented in form of a theorem. It takes into account very important fact that if all samples are available then the error will be zero, for both sparse and nonsparse recovered signals. Theory is illustrated and checked on statistical examples.

1. Introduction

A signal can be transformed from one domain into another in various ways. Some signals that cover the whole considered interval in one domain (where signals are dense in that domain) could be located within much smaller regions in another domain. We say that signals are sparse in a transformation domain if the number of nonzero coefficients is much smaller that the total number of signal samples. For example, a sum of discrete-time complex sinusoidal signals, with a number of components being much lower than the number of signal samples in the time domain, is a sparse signal in the discrete Fourier transform (DFT) domain.

Sparse signals could be reconstructed from much fewer samples than the sampling theorem

---

requires. Compressive sensing is a field dealing with the problem of signal recovery from a reduced set of samples [1]-[19]. As a study case, in this paper we will consider signals that are sparse in the Fourier transform domain. Signal sparsity in the discrete Fourier domain imposes some restrictions on the signal. Reducing the number of samples in the analysis manifests as a noise, whose properties are studied in [12] and used in [20] to define a reconstruction algorithm. The input noise influence is also an important topic in this analysis since the reduced number of available samples could increase the sensitivity of the recovery results to this noise [7, 20, 21]. Additive noise will remain in the resulting transform.

However, if a reconstruction algorithm for a sparse signal is used in the reconstruction of nonsparse signal then the noise, due to missing samples, will remain and behave as an additive input noise. A relation for the mean square error of this error is derived for the partial DFT matrix case. It takes into account very important fact that if all samples are available then the error will be zero, for both sparse and nonsparse recovered signals. Theory is illustrated and checked on statistical examples.

The paper is organised as follows: after the introduction part in Section 1, the definition of sparsity is presented in Section 2. In Section 3, the reconstruction algorithm is presented for both one step reconstruction and the iterative way. Also in Section 3, the analysis of the influence of additive noise will be expanded. The reconstruction of nonsparse signals with additive noise is shown in Section 4. In the appendix the conditions in which the reconstruction of sparse signals is possible in general are presented.

2. Sparsity and Reduced Set of Samples/Observations

Consider a signal $x(n)$ and its transformation domain coefficients $X(k)$,

$$x(n) = \sum_{k=0}^{N-1} X(k)\psi_k(n)$$

or

$$\mathbf{x} = \mathbf{\Psi X},$$

where $\mathbf{\Psi}$ is the transformation matrix with elements $\psi_k(n)$, $\mathbf{x}$ is the signal vector column, and $\mathbf{X}$ is the transformation coefficients vector column. For the DFT $\psi_k(n) = \exp(j2\pi nk)/N$. A signal is sparse in the transformation domain if the number of nonzero transform coefficients K is much lower than the number of the original signal samples N, Fig. 1, i.e., if

$$X(k) = 0$$

for

$$k \notin \{k_1, k_2, ..., k_K\} = \mathbf{K},$$

The number of nonzero samples is

$$\|\mathbf{X}\|_0 = \text{card}\{\mathbf{X}\} = K,$$

where

$$\|\mathbf{X}\|_0 = \sum_{k=0}^{N-1} |X(k)|^0$$

and card $\{\mathbf{X}\}$ is the notation for the number of nonzero transformation coefficients in $\mathbf{X}$. Counting the nonzero coefficients in a signal representation can be achieved by using the so called ℓ_0-norm denoted by $\|\mathbf{X}\|_0$. This form is referred to as the ℓ_0-norm (norm-zero) although it does not satisfy norm properties. By definition $|X(k)|^0 = 0$ for $|X(k)| = 0$ and $|X(k)|^0 = 1$ for $|X(k)| \neq 0$.

A signal $x(n)$, whose transformation coefficients are $X(k)$, is sparse in this transformation domain if

$$\text{card}\,\{\mathbf{X}\} = K \ll N.$$

For linear signal transforms the signal can be written as a linear combination of the sparse domain coefficients $X(k)$

$$x(n) = \sum_{k \in \{k_1, k_2, ..., k_K\}} X(k)\psi_k(n). \tag{1}$$

A signal sample can be considered as a linear combination (measurement) of values $X(k)$. Assume that samples of $x(n)$ are available only at a reduced set of random positions

$$n_i \in \mathbf{M} = \{n_1, n_2, ..., n_M\} \subset \mathbf{N} = \{0, 1, 2, 3, ..., N-1\}.$$

Here $\mathbf{N} = \{0, 1, 2, 3, ..., N-1\}$ is the set of all samples of a signal $x(n)$ and $\mathbf{M} = \{n_1, n_2, ..., n_M\}$ is its random subset with M elements, $M \leq N$. The available signal values are denoted by vector $\mathbf{y}$, Fig.1,

$$\mathbf{y} = [x(n_1),\ x(n_2),\ ..., x(n_M)]^T.$$

The available samples (measurements of a linear combination of $X(k)$) defined by (1), for $n_i \in \mathbf{M} = \{n_1, n_2, ..., n_M\}$, can be written as a system of M equations

$$\begin{bmatrix} x(n_1) \\ x(n_2) \\ ... \\ x(n_M) \end{bmatrix} = \begin{bmatrix} \psi_0(n_1) & \psi_1(n_1) & \psi_{N-1}(n_1) \\ \psi_0(n_2) & \psi_1(n_2) & \psi_{N-1}(n_2) \\ ... & ... & ... \\ \psi_0(n_M) & \psi_1(n_M) & \psi_{N-1}(n_M) \end{bmatrix} \begin{bmatrix} X(0) \\ X(0) \\ ... \\ X(N-1) \end{bmatrix}$$
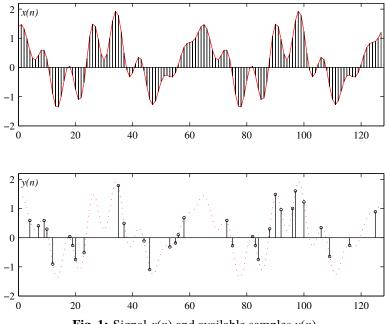
or

$$\mathbf{y} = \mathbf{A}\mathbf{X}$$

where $\mathbf{A}$ is the $M \times N$ matrix of measurements/observations/available signal samples.

The fact that the signal is sparse with $X(k) = 0$ for $k \notin \{k_1, k_2, ..., k_K\} = \mathbf{K}$ is not included in the measurement matrix $\mathbf{A}$ since the positions of the nonzero values are unknown. If the knowledge that $X(k) = 0$ for $k \notin \{k_1, k_2, ..., k_K\} = \mathbf{K}$ were included then a reduced observation matrix would be obtained as

$$\begin{bmatrix} x(n_1) \\ x(n_2) \\ ... \\ x(n_M) \end{bmatrix} = \begin{bmatrix} \psi_{k_1}(n_1) & \psi_{k_2}(n_1) & \psi_{k_K}(n_1) \\ \psi_{k_1}(n_2) & \psi_{k_2}(n_2) & \psi_{k_K}(n_2) \\ ... & ... & ... \\ \psi_{k_1}(n_M) & \psi_{k_2}(n_M) & \psi_{k_K}(n_M) \end{bmatrix} \begin{bmatrix} X(k_1) \\ X(k_2) \\ ... \\ X(k_K) \end{bmatrix}$$

or

$$\mathbf{y} = \mathbf{A}_K \mathbf{X}_K.$$

Fig. 1: Signal $x(n)$ and available samples $y(n)$.

Matrix $\mathbf{A}_K$ would be formed if we knew the positions of nonzero samples $k \in \{k_1, k_2, ..., k_K\} = \mathbf{K}$. It would follow from the measurement matrix $\mathbf{A}$ by omitting the columns corresponding to the zero-valued coefficients $X(k)$.

Assuming that there are K nonzero coefficients $X(k)$, out of the total number of N values, the total number of possible different matrices $\mathbf{A}_K$ is equal to the number of combinations with K out of N. It is equal to $\binom{N}{K}$.

3. Signal Reconstruction

Although the ℓ_0-norm cannot be used in the direct minimization, the algorithms based on the assumption that some coefficients $X(k)$ are equal to zero, and the minimization of the number of remaining nonzero coefficients that can reconstruct sparse signal, may efficiently be used.

A. Direct Combinatorial Search

The reconstruction process can be formulated as finding the positions and the values of K nonzero coefficients $X(k)$ of a sparse signal (or all signal $x(n)$ values) using a reduced set of signal values $x(n_i)$,

$$n_i \in \mathbf{M} = \{n_1, n_2, ..., n_M\} \subset \{0, 1, 2, ..., N - 1\}$$

such that

$$\min \|\mathbf{X}\|_0 \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{X}$$

where $\|\mathbf{X}\|_0 = \text{card}\{\mathbf{X}\} = K$. Consider a discrete-time signal $x(n)$. Signal is sparse in a transformation domain defined by the basis functions set $\psi_k(n)$, $k = 0, 1, ..., N-1$. The number of nonzero transform coefficients K is much lower than the number of the original signal samples N, i.e., $X(k) = 0$ for

$$k \notin \{k_1, k_2, ..., k_K\} = \mathbf{K},$$

$K \ll N$. A signal

$$x(n) = \sum_{k \in \{k_1, k_2, ..., k_K\}} X(k)\psi_k(n). \tag{2}$$

of sparsity K can be reconstructed from M samples, where $M \leq N$. In the case of signal $x(n)$ which is sparse in the transformation domain there are K nonzero unknown values $X(k_1)$, $X(k_2)$,...,$X(k_K)$. Other transform coefficients $X(k)$, for $k \notin \{k_1, k_2, ..., k_K\} = \mathbf{K}$, are zero-valued.

Just for the beginning assume that the transformation coefficient positions $\{k_1, k_2, ..., k_K\}$ are known. Then the minimal number of equations to find the unknown coefficients (and to calculate signal $x(n)$ for any n) is K. The equations are written for at least K time instants n_i, $i = 1, 2, ..., M \geq K$, where the signal is available/measured,

$$\sum_{k \in \mathbf{K}} X(k)\psi_k(n_i) = x(n_i), \text{ for } i = 1, 2, ..., M \geq K. \tag{3}$$

In a matrix form this system of equations is

$$\mathbf{A}_K \mathbf{X}_K = \mathbf{y}, \tag{4}$$

where $\mathbf{X}_K$ is the vector of unknown nonzero coefficients values (at the known positions) and $\mathbf{y}$ is the vector of available signal samples,

$$\mathbf{X}_K = [X(k_1) \ X(k_2) \ ... \ X(k_K)]^T \tag{5}$$

$$\mathbf{y} = [x(n_1) \ x(n_2) \ ... \ x(n_M)]^T$$

$$\mathbf{A}_K = \begin{bmatrix} \psi_{k_1}(n_1) & \psi_{k_2}(n_1) & ... & \psi_{k_K}(n_1) \\ \psi_{k_1}(n_2) & \psi_{k_2}(n_2) & ... & \psi_{k_K}(n_2) \\ ... & ... & ... & .... \\ \psi_{k_1}(n_K) & \psi_{k_2}(n_K) & ... & \psi_{k_K}(n_K) \end{bmatrix}. \tag{6}$$

Matrix $\mathbf{A}_K$ is the measurements matrix $\mathbf{A}$ with the columns corresponding to the zero-valued transform coefficients $k \notin \{k_1, k_2, ..., k_K\}$ being excluded. For a given set $\{k_1, k_2, ..., k_K\} = \mathbf{K}$ the coefficients reconstruction condition can be easily formulated as the condition that system (4) has a (unique) solution, i.e., that there are K independent equations,

$$\text{rank}(\mathbf{A}_K) = K.$$

Note that this condition does not guarantee that another set $\{k_1, k_2, ..., k_K\} = \mathbf{K}$ can also have a (unique) solution, for the same set of available samples. It requires $\text{rank}(\mathbf{A}_{2K}) = 2K$ for any submatrix $\mathbf{A}_{2K}$ of the measurements matrix $\mathbf{A}$.

System (3) is used with $K \ll M \leq N$. Its solution, in the mean squared sense, follows from the minimization of the difference of the available signal values and the values produced by

inverse transform of the reconstructed coefficients, $\min_{X(k)}\left\{e^2\right\}$ where

$$e^2 = \sum_{n\in\mathbf{M}}\left|y(n) - \sum_{k\in\mathbf{K}}X(k)\psi_k(n)\right|^2 =$$

$$= (\mathbf{y} - \mathbf{A}_K\mathbf{X}_K)^H(\mathbf{y} - \mathbf{A}_K\mathbf{X}_K) = \|\mathbf{y}\|_2^2 - 2\mathbf{X}_K^H\mathbf{A}_K^H\mathbf{y} + \mathbf{X}_K^H\mathbf{A}_K^H\mathbf{A}_K\mathbf{X}_K \qquad (7)$$

or

$$\min\left\{(\mathbf{y} - \mathbf{A}_K\mathbf{X}_K)^H(\mathbf{y} - \mathbf{A}_K\mathbf{X}_K)\right\}$$

where exponent H denotes the Hermitian conjugate. The derivative of e^2 over a specific co-efficient $X^*(p)$, $p \in \mathbf{K}$, is

$$\frac{\partial e^2}{\partial X^*(p)} = 2\left[\sum_{n\in\mathbf{M}}y(n) - \sum_{k\in\mathbf{K}}X(k)\psi_k(n)\right]\psi_p^*(n).$$

The minimum of quadratic form error is reached for $\partial e^2/\partial X^*(p) = 0$ when

$$\sum_{n\in\mathbf{M}}\psi_p^*(n)y(n) = \sum_{n\in\mathbf{M}}\sum_{k\in\mathbf{K}}\psi_k(n)\psi_p^*(n)X(k)$$

$$\text{for } p \in \mathbf{K}.$$

In matrix form this system of equations reads

$$\mathbf{A}_K^H\mathbf{y} = \mathbf{A}_K^H\mathbf{A}_K\mathbf{X}_K.$$

Its solution is

$$\mathbf{X}_K = \left(\mathbf{A}_K^H\mathbf{A}_K\right)^{-1}\mathbf{A}_K^H\mathbf{y}. \qquad (8)$$

It can be obtained by a symbolic vector derivation of (7) as

$$\frac{\partial e^2}{\partial \mathbf{X}_K^H} = -2\mathbf{A}_K^H\mathbf{y} + 2\mathbf{A}_K^H\mathbf{A}_K\mathbf{X}_K = 0.$$

If we do not know the positions of the nonzero values $X(k)$ for $k \in \{k_1, k_2, ..., k_K\} = \mathbf{K}$ then all possible combinations of $\{k_1, k_2, ..., k_K\} \subset \mathbf{N}$ should be tested. There are $\binom{N}{K}$ of them. It is not a computationally feasible problem. Thus we must try to find a method to estimate $\{k_1, k_2, ..., k_K\}$ in order to recover values of $X(k)$.

B. Estimation of Unknown Positions

Solution of the minimization problem, assuming that the positions of the nonzero signal coefficients in the sparse domain are known, is presented in the previous subsection. The next step is to estimate the coefficient positions, using the available samples. A simple way is to try to estimate the positions based on signal samples that are available, ignoring unavailable samples. This kind of transform estimate is

$$\hat{X}(k) = \sum_{n\in\mathbf{M}}x(n)\varphi_k(n), \qquad (9)$$

where for the DFT $\varphi_k(n) = \exp(-j2\pi nk/N)$ and $n \in \mathbf{M} = \{n_1, n_2, ..., n_M\}$. Since $\varphi_k(n) = N\psi_k^*(n)$ this relation can be written as

$$\hat{\mathbf{X}} = N\mathbf{A}^H\mathbf{y}$$

where $\mathbf{A}$ is the measurement matrix. With $K \ll M \ll N$ the coefficients $\hat{X}(k)$, calculated with M samples, are random variables. Note that using (9) in calculation is the same as assuming that the values of unavailable samples $x(n)$, $n \notin \mathbf{M}$, is zero. This kind of calculation corresponds to the result that would be achieved for the signal transform if ℓ_2-norm is used in minimization.

Algorithm

A simple and computationally efficient algorithm, for signal recovery, can now be implemented as follows:

(i) Calculate the initial transform estimate $\hat{X}(k)$ by using the available/remaining signal values

$$\hat{X}(k) = \sum_{n \in \mathbf{M}} x(n)\varphi_k(n) \tag{10}$$

or $\hat{\mathbf{X}} = N\mathbf{A}^H\mathbf{y}.$

(ii) Set the transform values $X(k)$ to zero at all positions k except the highest ones. Alternative:

(ii) Set the transform values $X(k)$ to zero at all positions k where this initial estimate $\hat{X}(k)$ is below a threshold T_r,

$$X(k) = 0 \ \text{ for } k \neq k_i, \ i = 1, 2, ..., \hat{K}$$
$$k_i = \arg\{|\hat{X}(k)| > T_r\}.$$

This criterion is not sensitive to T_r as far as all nonzero positions of the original transform are detected ($\hat{X}(k)$ is above the threshold) and the total number $\hat{K}$ of transform values in $\hat{X}(k)$ above the threshold is lower than the number of available samples, i.e., $K \leq \hat{K} \leq M$.

All $\hat{K} - K$ transform values that are zero in the original signal will be found as zero-valued.

(iii) The unknown nonzero (including $\hat{K} - K$ zero-valued) transform coefficients could be then easily calculated by solving the set of M equations for available instants $n \in \mathbf{M}$, at the detected nonzero candidate positions k_i, $i = 1, 2, ..., \hat{K}$,

$$\sum_{i=1}^{\hat{K}} X(k_i)\psi_{k_i}(n) = x(n), \ \text{ for } n \in \mathbf{M}.$$

This system of the form

$$\mathbf{A}_K\mathbf{X}_K = \mathbf{y}$$

is now reduced to the problem with known positions of nonzero coefficients (considered in the previous subsection). It is solved in the least square sense as (8)

$$\mathbf{X}_K = \left(\mathbf{A}_K^H\mathbf{A}_K\right)^{-1}\mathbf{A}_K^H\mathbf{y}. \tag{11}$$

The reconstructed coefficients $X(k_i)$, $i = 1, 2, ..., \hat{K}$, (denoted by vector $\mathbf{X}_K$) are exact, for all frequencies. If some transform coefficients, whose true value should be zero, are included (when $K < \hat{K}$) the resulting system will produce their correct (zero) values.

Comments: In general, a simple strategy can be used by assuming that $\hat{K} = M$ and by setting to zero value only the smallest $N - M$ transform coefficients in $\hat{X}(k)$. System (3) is then a system of M linear equations with $\hat{K} = M$ unknown transform values $X(k_i)$. If the algorithm fails to detect a component the procedure can be repeated after the detected components are reconstructed and removed. This simple strategy is very efficient if there is no input noise. Large $\hat{K}$, close or equal to M, will increase the probability that full signal recovery is achieved in one step. In this paper, it will be shown that in the case of an additive (even small) input noise in all signal samples, a reduction of the number $\hat{K}$ as close to the true signal sparsity K as possible will improve the signal to noise ratio.

Example: Consider a discrete signal

$$x(n) = 1.2e^{j2\pi n/16+j\pi/4} + 1.5e^{j14\pi n/16-j\pi/3} + 1.7e^{j12\pi n/16},$$

for $0 \leq n \leq 15$, sparse in the DFT domain since only three DFT values are different than zero. Assume now that its samples $x(2)$, $x(4)$, $x(11)$, and $x(14)$ are not available. We will show that, in this case, the exact DFT reconstruction may be achieved by:

(i) Calculating the initial DFT estimate by setting unavailable sample values to zero

$$\hat{X}(k) = \sum_{n\in\mathbf{M}} x(n)e^{j2\pi kn/16} = 16\mathbf{A}^H\mathbf{y},$$

where $n \in \mathbf{M} = \{0, 1, 3, 5, 6, 7, 8, 9, 10, 12, 13, 15\}$.

(ii) Detecting, for example $K = 3$ positions of maximal DFT values, k_1, k_2, and k_3, and (3) calculating the reconstructed DFT values at k_1, k_2, and k_3 from system

$$\frac{1}{16}\sum_{i=1}^{3} X(k_i)e^{j2\pi k_i n/16} = x(n),$$

where $n \in \mathbf{M} = \{0, 1, 3, 5, 6, 7, 8, 9, 10, 12, 13, 15\}$ are the instants where the signal is available.

The discrete-time signal $x(n)$, with $0 \leq n \leq 15$ is shown in Fig. 2. The signal is sparse in the DFT domain since only three DFT values are different than zero (Fig. 2 (second row)). The CS signal, with missing samples $x(2)$, $x(4)$, $x(11)$, and $x(14)$, being set to 0 for the initial DFT estimation, is shown in Fig. 2 (third row). The DFT of the signal, with missing values being set to 0, is calculated and presented in Fig. 2 (fourth row). There are three DFT values, at $k_1 = 1$, $k_2 = 6$, and $k_3 = 7$

$$\mathbf{K} = \{1, 6, 7\}$$

above the assumed threshold, for example, at level of 11. The rest of the DFT values is set to 0. This is justified by using the assumption that the signal is sparse. Now, we form a set of equations, for these frequencies $k_1 = 1$, $k_2 = 6$, and $k_3 = 7$ as

$$\frac{1}{16}\sum_{i=1}^{3} X(k_i)e^{j2\pi k_i n/16} = x(n),$$

where $n \in \mathbf{M} = \{0, 1, 3, 5, 6, 7, 8, 9, 10, 12, 13, 15\}$ are the instants where the signal is available. Since there are more equations than unknowns, the system $\mathbf{A}_K \mathbf{X}_K = \mathbf{y}$ is solved using $\mathbf{X}_K = \left(\mathbf{A}_K^H \mathbf{A}_K\right)^{-1} \mathbf{A}_K^H \mathbf{y}$. The obtained reconstructed values are exact, for all frequencies k, as in Fig. 2 (second row). They are shown in Fig. 2 (fifth row).

If the threshold was lower, for example at 7, then six DFT values at positions

$$\mathbf{K} = \{1, 6, 7, 12, 14, 15\}$$

are above the assumed threshold. The system with six unknowns

$$\frac{1}{16} \sum_{i=1}^{6} X(k_i) e^{j2\pi k_i n/16} = x(n),$$

where $n \in \mathbf{M} = \{0, 1, 3, 5, 6, 7, 8, 9, 10, 12, 13, 15\}$ will produce the same values for $X(1)$, $X(6)$, and $X(7)$ while the values $X(12) = X(14) = X(15) = 0$ will be obtained.

If the threshold is high to include the strongest signal component only, then the solution is obtained through an iterative procedure described in the next subsection.

C. Iterative Procedure

If components with very different amplitudes exist and the number of available samples is not large, then the iterative procedure should be used. This procedure could be implemented as follows. The largest component is detected and estimated first. It is subtracted from the signal. The next one is detected and the signal is estimated using the frequency from this and the previous step(s). The estimated two components are subtracted from the original signal. The frequency of next components is detected, and the process of estimations and subtractions is continued until the energy of the remaining signal is negligible or bellow an expected additive noise level.

Algorithm
(i) Calculate the initial transform estimate $\hat{X}_1(k)$ by using the available/remaining signal values $x_1(n) = x(n)$

$$\hat{X}_1(k) = \sum_{n \in \mathbf{M}} x(n)\varphi_k(n)$$

Set the transform values $\hat{X}(k)$ to zero at all positions k except the highest one at $k = k_1$, $\hat{\mathbf{K}}_1 = \{k_1\}$. Set the counter to $r = 1$.

Form the matrix $\mathbf{A}_1$ using the available samples in time $n \in \mathbf{N}_A$ and detected index $k \in \hat{\mathbf{K}}_1$, with one nonzero component. Calculate the estimate of the transformation coefficient at $k = k_1$
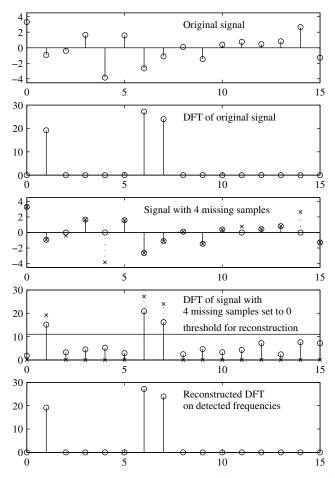
$$\hat{\mathbf{X}}_1 = \left(\mathbf{A}_1^H \mathbf{A}_1\right)^{-1} \mathbf{A}_1^H \mathbf{y}.$$

Calculate the signal estimation (as the inverse DFT)

$$\hat{x}_1(n) = \hat{X}_1(k_1)\psi_{k_1}(n), \text{ for } n \in \mathbf{M}$$

and check

$$\epsilon = \frac{\sum_{n \in \mathbf{M}} |x(n) - \hat{x}_1(n)|^2}{\sum_{n \in \mathbf{M}} |x(n)|^2}.$$

Fig. 2: Original signal in the discrete-time domain (first row); the DFT of the original signal (second row); signal with four missing samples at $n = 2, \ 4, 11,$ and 14 set to zero (third row); the DFT of signal with missing values being set to 0 (fourth row). The reconstructed signal assuming that the DFT contains components only at frequencies where the initial DFT is above threshold (fifth row). Absolute values of the DFT and real part of signal are shown.

If, for example $\epsilon < 10^{-5}$, stop the calculation and use $x(n) = \hat{x}_1(n)$. If not then go to the next step.

(ii) Set the counter to $r = r + 1$. Form a signal

$$e_r(n) = x(n) - \hat{x}_{r-1}(n),$$

at the available sample positions and calculate the transform

$$\hat{E}_r(k) = \sum_{n \in \mathbf{M}} e_r(n)\varphi_k(n).$$

Set the transform values $\hat{E}_r(k)$ to zero at all positions k except the highest one at $k = k_r$. Form the set of r indices, using union of the previous maxima positions and the detected position, as

$$\hat{\mathbf{K}}_r = \{\hat{\mathbf{K}}_{r-1}, k_r\}.$$

Form matrix $\mathbf{A}_r$ using the available samples in time $n \in \mathbf{M}$ and detected $\hat{K}_r$ indices $k \in \hat{\mathbf{K}}_r$. Calculate the estimate of K_r transformation coefficients

$$\hat{\mathbf{X}}_{K_r} = \left(\mathbf{A}_r^H \mathbf{A}_r\right)^{-1} \mathbf{A}_r^H \mathbf{y}.$$

Calculate the signal

$$\hat{x}_r(n) = \sum_{i=1}^{\hat{K}_r} \hat{X}_r(k_i)\psi_{k_i}(n), \text{ for } n \in \mathbf{M}$$

and check

$$\epsilon = \frac{\sum_{n \in \mathbf{M}} |x(n) - \hat{x}_r(n)|^2}{\sum_{n \in \mathbf{M}} |x(n)|^2}.$$

If, for example $\epsilon < 10^{-5}$, stop the calculation and use
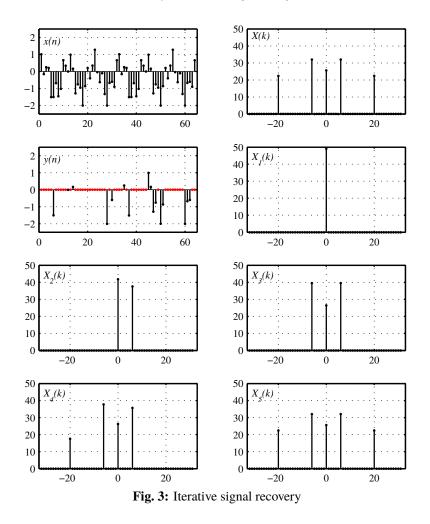
$$x(n) = \hat{x}_r(n).$$

Else repeat step (ii).
   Example: Signal

$$x(n) = \sin(12\pi \frac{n}{N} + \frac{\pi}{4}) + 0.7 \cos(40\pi \frac{n}{N} + \frac{\pi}{3}) - 0.4$$

with $N = 64$ is shown in Fig.3. Small number of samples is available $M = 16$ with different signal amplitudes, making one-step recovery impossible. The available signal samples $y(n)$ are shown in Fig.3 (second row, left). The iterative procedure is used and, for the detected DFT positions during the iterations, the recovered signal is calculated according to the presented algorithm. The recovered DFT values in the rth iteration are denoted as $X_r(k)$ and presented in Fig.3. After first iteration the strongest component is detected and its amplitude is estimated. At this stage, other components behave as noise and make amplitude value inaccurate. Accuracy improves as the number of detected components increases in next iterations. After five steps the agreement between the reconstructed signal and the available signal samples was complete. Then the algorithm is stopped. The DFT of the recovered signal is presented as $X_5(k)$ in the last subplot of Fig.3. Its agreement with the DFT of the original signal, Fig.3 (first row, right) is complete.

D. Unavailable/Missing Samples Noise

The initial DFT calculation in reconstruction algorithms is done assuming zero-valued missing samples. The initial calculation quality has a crucial importance to the successful signal recovery. With a large number of randomly positioned missing samples, the missing

Fig. 3: Iterative signal recovery

samples manifest as a noise in this initial transform. Once the reconstruction conditions are met for a sparse signal and the exact reconstruction is achieved, the noise due to missing samples does not influence the results in a direct way. It influences the possibility to recover a signal. The accuracy of the recovery results is related to the additive input noise only. The input noise is transformed by the recovery algorithm into a new noise depending on the signal sparsity and the number of available samples. A simple analysis of this form of noise will be presented in the second part of this section.

Consider a sparse signal in the DFT domain with nonzero coefficients $X(k)$ at the positions $k \in \mathbf{K} = \{k_1, k_2, ..., k_K\}$

$$x(n) = \sum_{p=1}^{K} A_p e^{j2\pi n k_p / N},$$

where A_p are the signal component amplitudes. The initial DFT is calculated using $n \in \mathbf{M} = \{n_1, n_2, ..., n_M\}$

$$X(k) = \sum_{n \in \mathbf{M}} x(n) e^{-j2\pi nk/N} = \sum_{n \in \mathbf{M}} \sum_{p=1}^{K} A_p e^{-j2\pi n(k-k_p)/N}. \tag{12}$$

We can distinguish two cases:

(1) For $k = k_i \in \{k_1, k_2, ..., k_K\}$ then, with $M = \text{card}(\mathbf{M})$,

$$X(k_i) = A_i M + \sum_{n \in \mathbf{M}} \sum_{\substack{p=1 \\ p \neq i}}^{K} A_p e^{-j2\pi n(k_i - k_p)/N}.$$

The value of

$$\Xi = \sum_{n \in \mathbf{M}} \sum_{\substack{p=1 \\ p \neq i}}^{K} A_p e^{-j2\pi n(k_i - k_p)/N} \tag{13}$$

with random set $\mathbf{M} = \{n_1, n_2, ..., n_M\}$, for $1 \ll M \ll N$, can be considered as a random variable. Its mean over different realizations of available samples (different realizations of sets $\mathbf{M}$) is $\mathrm{E}\{\Xi\} = 0$. The mean value of $X(k_i)$ is

$$\mathrm{E}\{X(k_i)\} = A_i M.$$

(2) For $k \notin \{k_1, k_2, ..., k_K\}$ the mean value of (12) is

$$\mathrm{E}\{X(k)\} = 0.$$

The mean value of (12) for any k is of the form

$$\mathrm{E}\{X(k)\} = M \sum_{p=1}^{K} A_p \delta(k - k_p),$$

Its variance is [12, 23]

$$\sigma_N^2(k) = \mathrm{var}(X(k)) = \sum_{p=1}^{K} A_p^2 M \frac{N-M}{N-1} \left[1 - \delta(k - k_p)\right]. \tag{14}$$

The ratio of the signal amplitude $X(k_1)$ and the standard deviation $\sigma_N(k)$ for $k \neq k_1$ is crucial parameter (Welsh bound for coherence index μ of measurement matrix $\mathbf{A}$) for correct signal detection. Its value is

$$\frac{\sigma_N(k)}{|X(k_1)|} = \sqrt{\frac{N-M}{M(N-1)}}.$$

Note that the variance in a multicomponent signal with $K > 1$ is a sum of the variances of individual components at all frequencies k except at $k_i \in \{k_1, k_2, ..., k_K\}$ when the values are lower for $|A_i|^2 M \frac{N-M}{N-1}$ since all component values are added up in phase, without random variations.

According to the central limit theorem, for $1 \ll M \ll N$ the real and imaginary parts of the DFT value for noise only positions $k \notin \{k_1, k_2, ..., k_K\}$ can be described by Gaussian distribution, $\mathcal{N}(0, \sigma_N^2/2)$ with zero-mean and variance $\sigma_N^2 = \sigma_N^2(k)$. Real and imaginary parts of the DFT value, at the signal component position $k_i \in \{k_1, k_2, ..., k_K\}$, can be described by the Gaussian distributions

$$\mathcal{N}(M \operatorname{Re}\{A_p\}, \sigma_{S_p}^2/2), \text{ and}$$

$$\mathcal{N}(M \operatorname{Im}\{A_p\}, \sigma_{S_p}^2/2),$$

respectively, where $\sigma_{S_p}^2 = \sigma_N^2 - A_p^2 M \frac{N-M}{N-1}$, according to (14).

Example: For a discrete-time signal

$$x(n) = A_1 e^{j2\pi k_1 n/N} + A_2 e^{j2\pi k_2 n/N} + A_2 e^{j2\pi k_3 n/N}, \tag{15}$$

with $N = 64, A_1 = 1, A_2 = 1/2, A_3 = 1/4$, the DFT is calculated using a random set of $M = 16$ samples. Calculation is performed with 10^5 random realizations with randomly positioned M samples and random values of $k_1, k_2,$ and k_3. Histogram of the DFT values, at a noise only position $k \notin \{k_1, k_2, k_3\}$ and at the signal component $k = k_1$ position, is presented in Fig.4 (left). Histogram of the DFT real part is shown, along with the corresponding Gaussian functions $\mathcal{N}(0, \frac{21}{16} \frac{N-M}{N-1})$ and $\mathcal{N}(M, \frac{5}{16} \frac{N-M}{N-1})$, shown by green dots. The same calculation is repeated with $M = 64$, Fig.4 (right). We can see that the mean value of the Gaussian variable $X(k)$ can be used for the detection of the signal component position. Also the variance is different for noise only and the signal component positions. It can also be used for the signal position detection. In the case with $M = 16$, the histograms are close to each other, meaning that there is a probability that a signal component is missdetected. Histograms are well separated in the case when $M = 64$. It means that the signal component will be detected with an extremely high probability in this case. Calculation of the detection probability is straightforward with the assumed probability density functions.

The spark based relation can be obtained within the framework of the previous analysis if we assume that the noises (13) due to missing samples coming from different components of the same (unity) amplitude A_i are added up (equal amplitudes are the worst case for this kind of analysis) with the same phase to produce [23],

$$X(k) = \sum_{n \in \mathbf{M}} \sum_{p=1}^{K} e^{-j2\pi n(k-k_p)/N} = K \sum_{n \in \mathbf{M}} e^{-j2\pi n(k-k_p)/N} \tag{16}$$

at some frequency $k \notin \{k_1, k_2, ..., k_K\}$. Random variable $\sum_{n \in \mathbf{M}} e^{-j2\pi n(k-k_p)/N}$ (since $n \in \mathbf{M}$ is random) should also assume its maximal possible value (calculated over all possible k_p and all possible positions $k, k \neq k_p$). The maximal possible value of this variable is related to the coherence index μ of the partial DFT matrix defined as

$$\mu = \max |\mu(k, k_p)| = \frac{1}{M} \max_{k,k_p} \left| \sum_{n \in \mathbf{M}} e^{-j2\pi n(k-k_p)/N} \right|. \tag{17}$$

It means that maximal possible value of this variable is μM. It should also be assumed that $(K-1)$ remaining noise components (due to missing samples) at the component position
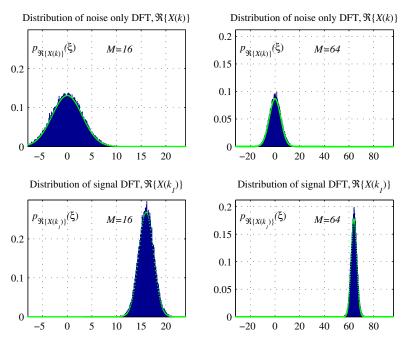
Fig. 4: Histograms and Gaussian probability density functions for the signal and noise only positions in the DFT for a three-component signal with $N = 128$ and $M = 16$ (left) and $M = 64$ (right). The histograms are calculated in 10^5 random realizations of M available samples and random signal frequency positions.

$k = k_p$ assume the same maximal value μM and that all of them subtract in phase from the signal mean value M at $k = k_p$. Condition for the correct detection of a component position at $k = k_p$ is then such that the minimal possible amplitude of the component $M - M\mu(K - 1)$ is greater than the maximal possible noise $M\mu K$ at $k \notin \{k_1, k_2, ..., k_K\}$, i.e.,

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

or

$$K < \frac{1}{2}(1 + \frac{1}{\mu}) = \frac{1}{2}\text{spark}(\mathbf{A}),$$

where spark($\mathbf{A}$) is the spark of the measurement matrix $\mathbf{A}$ (spark of matrix $\mathbf{A}$ is defined as the smallest number of dependent columns or rows). According to many very unlikely assumptions that has been made, we can state that this is a very pessimistic bound for K. Therefore, for a high degree of randomness, a probabilistic approach may be more suitable for the analysis than the spark based relation.

E. Additive Noise Influence

Assume an additive noise $\varepsilon(t)$ in the input signal. In a matrix form this system of M linear equations with K unknowns reads

$$\mathbf{y} + \varepsilon = \mathbf{A}\mathbf{X}_K$$

The solution follows for

$$\mathbf{X}_K = \mathbf{X}_{KS} + \mathbf{X}_{KN}$$

where $\mathbf{X}_{KS}$ and $\mathbf{X}_{KN}$ are the reconstructed signal and noise components respectively.

Assume that the reconstruction conditions are met and the positions of nonzero coefficients $\mathbf{K} = \{k_1, k_2, ..., k_K\}$ can be determined through a single step or iterative procedure [13, 20, 22]. The equations to find the unknown coefficients are written for $M > K$ time instants n_i, $i = 1, 2, ..., M \geq K$

$$\mathbf{A}_K\mathbf{X}_K = \mathbf{y} + \varepsilon \tag{18}$$

where $\mathbf{X}_K = [X(k_1) \; X(k_2) \; ... \; X(k_K)]^T$ is the vector of unknown nonzero coefficients values (at the determined positions) and $\mathbf{y}$ is the vector of the available signal samples $\mathbf{y} = [x(n_1) \; x(n_2) \; ... \; x(n_M)]^T$. The matrix $\mathbf{A}_K$ is the measurements matrix $\mathbf{A}$ with the columns corresponding to the zero-valued transform coefficients $k \notin \{k_1, k_2, ..., k_K\}$ being excluded. For a given set $\{k_1, k_2, ..., k_K\} = \mathbf{K}$ the coefficients reconstruction condition can be easily calculated as

$$\mathbf{X}_K = \left(\mathbf{A}_K^H\mathbf{A}_K\right)^{-1} \mathbf{A}_K^H(\mathbf{y} + \varepsilon). \tag{19}$$

where $\mathbf{X}_{KS} = \left(\mathbf{A}_K^H\mathbf{A}_K\right)^{-1} \mathbf{A}_K^H\mathbf{y}$ and $\mathbf{X}_{KN} = \left(\mathbf{A}_K^H\mathbf{A}_K\right)^{-1} \mathbf{A}_K^H\varepsilon$ is the noise influence to the reconstructed signal coefficients.

The input signal-to-noise (SNR) ratio, if all signal samples were available, is

$$SNR_i = 10\log \frac{\sum_{n=0}^{N-1} |x(n)|^2}{\sum_{n=0}^{N-1} |\varepsilon(n)|^2} = 10\log \frac{E_x}{E_\varepsilon}.$$

Assume the noise energy in M available samples used in the reconstruction is

$$E_{\varepsilon A} = \sum_{n \in \mathbf{M}} |\varepsilon(n)|^2 . \tag{20}$$

The correct amplitude in the signal transform at the frequency k_p, in the case if all signal samples were used, would be NA_p. To compensate the resulting transform for the known bias in amplitude when only M available samples are used we should multiply the coefficient by N/M. It means that is a full recovery, a signal transform coefficient should correspond to the coefficient of the original signal with all signal samples being used. The noise in the transform coefficients will also be multiplied by the same factor. Therefore, its energy would be increased to $E_{\varepsilon A}N^2/M^2$. The signal-to-noise ratio in the recovered signal would be

$$SNR = 10\log \frac{\sum_{n=0}^{N-1} |x(n)|^2}{\frac{N^2}{M^2} \sum_{n \in \mathbf{M}} |\varepsilon(n)|^2} \tag{21}$$

If the distribution of noise in the samples used for reconstruction is the same as in other signal samples then $\sum_{n \in \mathbf{M}} |\varepsilon(n)|^2 = M\sigma_\varepsilon^2$ and

$$SNR = 10\log \frac{\sum_{n=0}^{N-1} |x(n)|^2}{\frac{N^2}{M^2} M\sigma_\varepsilon^2} = SNR_i - 10\log\left(\frac{N}{M}\right). \tag{22}$$

Table I

Signal to noise ratio in the reconstructed signal according to the theory SNR_T and the statistics SNR_S for various M. Input signal to noise ratio is denoted by SNR_i

SNR in [dB]	$M = 128$	$M = 160$	$M = 192$	$M = 224$
SNR_i	3.5360	3.5326	3.5788	3.5385
SNR_T	18.5953	19.5644	20.3562	21.0257
SNR_S	18.7203	19.5139	20.2869	21.7302

Therefore, a signal reconstruction that would be based on the initial estimate (10) would worsen SNR, since $N > M$.

Since only K out of N DFT coefficients are used in the reconstruction the energy of the reconstruction error is reduced for the factor of K/N as well. Therefore, the energy of noise in the reconstructed signal is

$$E_{\varepsilon R} = \frac{K}{N} \frac{N^2}{M^2} \sum_{n \in \mathbf{M}} |\varepsilon(n)|^2 .$$

The output signal to noise ratio in the reconstructed signal is [20, 22, 23]

$$SNR = 10 \log \frac{\sum_{n=0}^{N-1} |x(n)|^2}{\frac{KN}{M^2} \sum_{n \in \mathbf{M}} |\varepsilon(n)|^2} = 10 \log \frac{\sum_{n=0}^{N-1} |x(n)|^2}{\frac{K}{M} \sum_{n=0}^{N-1} |\varepsilon(n)|^2}. \tag{23}$$

It is related to the input signal to noise ration SNR_i as

$$SNR = SNR_i - 10 \log \left(\frac{K}{M} \right). \tag{24}$$

Example: Theory is illustrated on a four component noisy signal

$$x(n) = A_1 \exp(j2\pi k_1 n/N) + A_2 \exp(j2\pi k_2 n/N)$$
$$+ A_3 \exp(j2\pi k_3 n/N) + A_4 \exp(j2\pi k_4 n/N) + \varepsilon(n)$$

as well, where $A_1 = 1$, $A_2 = 0.75$, $A_3 = 0.5$, $A_4 = 0.67$, $N = 257$, and $\{k_1, k_2, k_3, k_4\} = \{58, 117, 21, 45\}$. The signal is reconstructed using iterative calculation to find nonzero coefficients $\mathbf{K} = \{k_1, k_2, ..., k_K\}$ and (19) to find the signal. The results are presented in the Table I. The agreement of the numerical statistical results SNR_S with this simple theory in analysis of noise influence to the reconstruction of sparse signals SNR_T is high for all considered SNR_i.

4. Nonsparse Signal Reconstruction

According to the results in previous section, the missing samples can be represented by a noise influence. Assume that we use a reconstruction algorithm for a signal of sparsity K on a signal whose DFT coefficients $\mathbf{X}$ are not sparse (or not sufficiently sparse). Denote by $\mathbf{X}_K$ the sparse signal with K nonzero coefficients equal to the largest K coefficients of $\mathbf{X}$. Suppose that the number of components K and the measurements matrix satisfy the reconstruction conditions so that a reconstruction algorithm can detect (one by one or at once) largest K

components $(A_1, A_2, ... A_K)$ and perform signal reconstruction to get $\mathbf{X}_R$. The remaining $N - K$ components $(A_{K+1}, A_{K+2}, ..., A_N)$ will be treated as noise in these K largest components. Variance of a signal component is $|A_i|^2 M(N - M)/(N - 1)$. After reconstruction this variance is

$$|A_i|^2 \frac{N^2}{M^2} \frac{M(N - M)}{N - 1} \cong |A_i|^2 N \frac{N - M}{M}.$$

The total energy of noise in the reconstructed K largest components $\mathbf{X}_R$ will be

$$\|\mathbf{X}_R - \mathbf{X}_K\|_2^2 = KN \frac{N - M}{M} \sum_{i=K+1}^{N} |A_i|^2$$

Denoting the energy of remaining signal, when the K largest are removed from the original signal, by

$$\|\mathbf{X} - \mathbf{X}_K\|_2^2 = N \sum_{i=K+1}^{N} |A_i|^2$$

we get

$$\|\mathbf{X}_R - \mathbf{X}_K\|_2^2 = K \frac{N - M}{M} \|\mathbf{X} - \mathbf{X}_K\|_2^2 .$$

If the signal is sparse, i.e., $\mathbf{X} = \mathbf{X}_K$, then

$$\|\mathbf{X}_R - \mathbf{X}_K\|_2^2 = 0.$$

The same result follows if $N = M$

$$\|\mathbf{X}_R - \mathbf{X}_K\|_2^2 = 0.$$

That is, the error will be zero if a complete DFT matrix is used in calculation of any signal component.

Using Schwartz inequality $\|\mathbf{X}\|_2 \leq \frac{1}{\sqrt{N}} \|\mathbf{X}\|_1$ follows

$$\|\mathbf{X}_K - \mathbf{X}_R\|_2 \leq \sqrt{\frac{N - M}{M} \frac{K}{N - K}} \|\mathbf{X} - \mathbf{X}_K\|_1 .$$

It means that if $\|\mathbf{X} - \mathbf{X}_K\|_1$ is minimized then the upper bound of the error $\|\mathbf{X}_K - \mathbf{X}_K\|_2$ is also minimized.

Based on the previous results we can easily get the following result.

A. Theorem for the Error in a Nonsparse Signal Reconstruction

Theorem 1 *Consider a signal $x(n)$, $n = 0, 1, ..., N - 1$, with transformation coefficients $\mathbf{X}$ and unknown sparsity in this transformation domain, including the case when the signal is not sparse. Assume that the signal is reconstructed using the reconstruction algorithms assuming that its sparsity were K. Denote the reconstructed signal by $\mathbf{X_R}$ and the original signal transform at the nonzero value positions $k \in \mathbf{K}$ of the reconstructed signal coefficients by $\mathbf{X_K}$, where $\mathbf{X_K} = \mathbf{X}$ for $k \in \mathbf{K}$ and $\mathbf{X_K} = \mathbf{0}$ for $k \notin \mathbf{K}$. Assume that in $M \leq N$ available samples of the input signal in the time domain there is an additive white noise with variance σ_ε^2. The*

total error in the reconstructed signal, with respect to the original signal at the same nonzero coefficient positions, is

$$\|\mathbf{X}_K - \mathbf{X}_R\|_2^2 = K \frac{N-M}{M} \|\mathbf{X} - \mathbf{X}_K\|_2^2 + \frac{K}{M} N \sigma_\varepsilon^2.$$

Proof of this theorem easily follows from the presented analysis.

Example: Consider a nonsparse signal

$$x(n) = e^{j2\pi k_1 n/N} + 0.8 e^{j2\pi k_2 n/N} + 0.77 e^{j2\pi k_3 n/N} + 0.75 e^{j2\pi k_4 n/N}$$
$$+ \sum_{i=0}^{250} \left(\frac{1}{3}\right)^{1+i/50} e^{j2\pi k_{i+5} n/N}$$

where k_i, $i = 1, 2, ..., 251+4$ are random frequency indices from 0 to $N-1$. Using $N = 257$ and $M = 192$, the first $K = 4$ components of signal are reconstructed. The remaining 251 signal components are considered as disturbance. Reconstruction of $K = 4$ largest components is done in 100 independent realizations with different frequencies and positions of available samples. The result for SNR in the noise free case, obtained statistically and by using the Theorem, is

$$SNR_{stat} = 10 \log \left(\frac{\|\mathbf{X}_K\|_2^2}{\|\mathbf{X}_K - \mathbf{X}_R\|_2^2} \right) = 23.1476$$

$$SNR_{theor} = 10 \log \left(\frac{\|\mathbf{X}_K\|_2^2}{K \frac{N-M}{M} \|\mathbf{X} - \mathbf{X}_K\|_2^2} \right) = 23.1235.$$

Note that the calculation of SNR_{theor} is simple since we assumed that the amplitudes of disturbing components are coefficients of a geometric series. One realization with $K = 4$ is presented in Fig. 5. The case when $K = 10$ is presented in Fig. 6. Red signal (with dots) represents the reconstructed signal with assumed sparsity and the signal with black crosses represents the original nonsparse signal.

In the case of additive complex-valued noise of variance $\sigma_\varepsilon^2 = 2$ the results are

$$SNR_{stat} = 10 \log \left(\frac{\|\mathbf{X}_K\|_2^2}{\|\mathbf{X}_K - \mathbf{X}_R\|_2^2} \right) = 17.0593$$

$$SNR_{theor} = 10 \log \left(\frac{\|\mathbf{X}_K\|_2^2}{K \frac{N-M}{M} \|\mathbf{X} - \mathbf{X}_K\|_2^2 + \frac{K}{M} N \sigma_\varepsilon^2} \right) = 17.0384.$$

The decrease in the SNR due to noise is

$$\Delta SNR_{theor} = 10 \log \left(\frac{K \frac{N-M}{M} \|\mathbf{X} - \mathbf{X}_K\|_2^2}{K \frac{N-M}{M} \|\mathbf{X} - \mathbf{X}_K\|_2^2 + \frac{K}{M} N \sigma_\varepsilon^2} \right) = -6.0851.$$

The simulation is repeated with $M = 128$ and the same noise. The SNR values are $SNR_{theor} = 14.3345$ and $SNR_{stat} = 14.4980$.
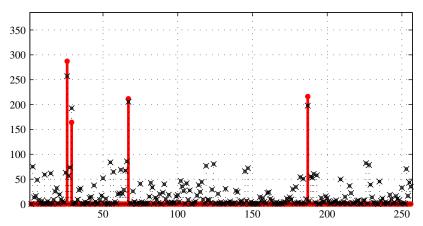
Fig. 5: Single realization reconstruction of $K = 4$ largest signal components of a nonsparse noisy signal.
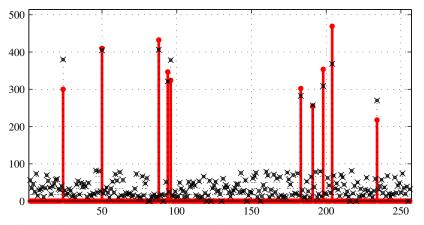

Fig. 6: Single realization reconstruction of $K = 10$ largest signal components of a nonsparse noisy signal.

5. Conclusions

The goal of compressive sensing is to reconstruct a sparse signal using a reduced set of available samples. It can be done by minimizing the sparsity measure and available samples. A simple algorithm for signal reconstruction is presented. One step reconstruction and an iterative procedure of the reconstruction algorithm are given. Noisy environment is taken into account as well. The input noise can degrade the reconstruction limit. However, as far as the reconstruction is possible, the noise caused by missing samples manifests its influence to the results accuracy in simple and direct way through the number of missing samples and

signal sparsity. The accuracy of the final result is related to the input noise intensity, number of available samples and the signal sparsity. A theorem presenting error in the case when the reconstruction algorithm defined for reconstruction of sparse signals are used in for nonsparse signals reconstruction is defined as well. The theory is checked and illustrated on numerical examples.

6. Appendix: Reconstruction Conditions

Consider an N-dimensional vector $\mathbf{X}$ whose sparsity is K and its M measurements $\mathbf{y} = \mathbf{AX}$, where the measurements matrix $\mathbf{A}$ is an $M \times N$ matrix, with $K < M \leq N$. A reconstruction vector $\mathbf{X}$ can be achieved from a reduced set of samples/measurements using the sparsity measures minimization.

The ℓ_1-norm based solution of constrained sparsity measure minimization

$$\min \|\mathbf{X}\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{AX} \tag{25}$$

is the same as the ℓ_0-norm based solution of

$$\min \|\mathbf{X}\|_0 \quad \text{subject to} \quad \mathbf{y} = \mathbf{AX}$$

if the measurements matrix $\mathbf{A}$ satisfies the restricted isometry property for a $2K$ sparse vector

$$(1 - \delta_{2K}) \|\mathbf{X}_{2K}\|_2^2 \leq \frac{1}{E_\Psi} \|\mathbf{A}_{2K}\mathbf{X}_{2K}\|_2^2 \leq (1 + \delta_{2K}) \|\mathbf{X}_{2K}\|_2^2$$

with a sufficiently small δ_{2K}. Constant E_Ψ is the energy of columns of measurement matrix $\mathbf{A}$. For normalized energy $E_\Psi = 1$, while for the measurement matrix obtained using M rows of the standard DFT matrix $E_\Psi = M$. If the signal $\mathbf{X}$ is not sparse then the solution of minimization problem (25) denoted by $\mathbf{X}_R$ will satisfy

$$\|\mathbf{X}_R - \mathbf{X}\|_2 \leq C_0 \frac{\|\mathbf{X}_K - \mathbf{X}\|_1}{\sqrt{K}} \tag{26}$$

where $\mathbf{X}_K$ is K sparse signal corresponding to K largest values of $\mathbf{X}$. If the signal $\mathbf{X}$ is of sparsity K then $\|\mathbf{X}_K - \mathbf{X}\|_2 = 0$ and $\mathbf{X}_R = \mathbf{X}$. In the case of noisy measurements when

$$\|\mathbf{y} - \mathbf{AX}\|_2 \leq \epsilon$$

then [3]

$$\|\mathbf{X}_R - \mathbf{X}\|_2 \leq C_0 \frac{\|\mathbf{X}_K - \mathbf{X}\|_1}{\sqrt{K}} + C_1 \epsilon$$

where C_0 and C_1 are constants depending on δ_{2K}. For example, with $\delta_{2K} = 1/4$ constants are $C_0 \leq 5.5$ and $C_1 \leq 6$, [3].

References

[1] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

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

[3] E. J. Candès, M. B. Wakin, "An Introduction to Compressive Sampling", *IEEE Signal Processing Magazine,* vol.21, March 2008.

[4] R. E. Carrillo, K. E. Barner, and T. C. Aysal, "Robust sampling and reconstruction methods for sparse signals in the presence of impulsive noise," *IEEE Journal of Selected Topics in Signal Processing*, 2010, 4(2), pp. 392–408.

[5] L. Stanković, M. Daković, and T. Thayaparan, *Time–Frequency Signal Analysis with Application*, Artech House, 2013.

[6] M. A. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems,"*IEEE Journal of Selected Topics in Signal Processing,*, vol. 1, no. 4, pp. 586–597, 2007.

[7] D. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on Information Theory*, vol. 52, pp. 6–18, 2006.

[8] B. Turlach, "On algorithms for solving least squares problems under an L1 penalty or an L1 constraint," *Proc. of the American Statistical Association; Statistical Computing Section*, pp. 2572–2577, Alexandria, VA, 2005.

[9] R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, 2007, pp. 118–121.

[10] P. Flandrin and P. Borgnat, "Time-Frequency Energy Distributions Meet Compressed Sensing," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, 2010, pp. 2974–2982.

[11] Y. D. Zhang and M. G. Amin, "Compressive sensing in nonstationary array processing using bilinear transforms," in Proc. *IEEE Sensor Array and Multichannel Signal Processing Workshop*, Hoboken, NJ, June 2012.

[12] L. Stanković, S. Stanković, and M. G. Amin, "Missing Samples Analysis in Signals for Applications to L-Estimation and Compressive Sensing", *Signal Processing*, Elsevier, Volume 94, Jan. 2014, Pages 401–408.

[13] L. Stanković, "Noises in Randomly Sampled Sparse Signals," Facta Universitatis, Series: Electronics and Energetics, Vol 27, No 3, September 2014.

[14] L. Stanković, "A measure of some time–frequency distributions concentration," *Signal Processing*, vol. 81, pp. 621–631, 2001

[15] E. Sejdić, A. Cam, L. F. Chaparro, C. M. Steele, and T. Chau, "Compressive sampling of swallowing accelerometry signals using TF dictionaries based on modulated discrete prolate spheroidal sequences," *EURASIP Journal on Advances in Signal Processing*, 2012:101 doi:10.1186/1687–6180–2012–101.

[16] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on pure and applied mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.

[17] L. Stanković, M. Daković, S. Vujović, "Adaptive variable step algorithm for missing samples recovery in sparse signals," *IET Signal Processing*, vol.8, no.3, 2014, pp.246–256, doi: 10.1049/iet-spr.2013.0385.

[18] I. Stanković, "Sparse Aignals: Analysis and Reconstruction", Final year project (project advisor D. Barjamovic), University of Westminster, June 2014.

[19] I. Stanković, "Track Global Ozone Density with Missing Data", MSc Thesis (project advisor Dr W. Dai), Imperial College London, September 2015.

[20] S. Stanković, I. Orović, and L. Stanković, "An Automated Signal Reconstruction Method based on Analysis of Compressive Sensed Signals in Noisy Environment", *Signal Processing*, Elsevier, Volume 94, in print.

[21] S. Stanković, I. Orović, and E. Sejdić, Multimedia Signals and Systems: Basic and Advance Algorithms for Signal Processing,. Springer-Verlag, New York, 2015.

[22] L. Stanković, "On the ISAR Image Analysis and Recovery with Unavailable or Heavily Corrupted Data", *IEEE Trans. Aerospace and Electronic Systems*, Vol.51, July 2015.

[23] L. Stanković, "*Digital Signal Processing*", CreateSpace Independent Publishing Platform (an Amazon.com Copmany), Nov. 2015.