

On the Uniqueness of the Sparse Signals Reconstruction Based on the Missing Samples Variation Analysis

Ljubiša Stanković, Miloš Daković

Abstract—An approach to sparse signals reconstruction considering its missing measurements/samples as variables is recently proposed. Number and positions of missing samples determine the uniqueness of the solution. It has been assumed that analyzed signals are sparse in the discrete Fourier transform (DFT) domain. A theorem for simple uniqueness check is proposed. Two forms of the theorem are presented, for an arbitrary sparse signal and for an already reconstructed signal. The results are demonstrated on illustrative and statistical examples.

I. INTRODUCTION

In many engineering applications, an incomplete set of samples/measurements arises due to physical system constraints. In some cases randomly positioned samples/measurements are heavily corrupted so that it is better to omit them and consider as unavailable [1], [2]. In these applications reduction of the considered dataset is not a result of an intentional compressive sensing strategy [3]-[8]. Nevertheless, the primary goal is the signal recovery, as in the compressive sensing theory [8], [16]. Recently, an adaptive-step gradient-based method for the reconstruction of sparse signals with missing/omitted samples has been proposed [17]. The proposed method reconstructs the remaining missing samples/measurements in order to make a complete set of samples/measurements, in contrast to the common reconstruction methods that recover the signal in their sparsity domain. The final result in all algorithms is the same. It is full recovery of the signal.

In general, the reconstructed signal uniqueness is guaranteed if the restricted isometry property is used and checked with appropriate isometry constants [5]. However, two problems exist in the implementation of this method. For a specific measurement matrix, it produces quite conservative bounds. In practice, it produces a large number of false alarms for nonuniqueness. In addition, uniqueness check with the restricted isometry property requires a combinatorial approach, which is an NP-hard problem (like the solution of the problem itself using the norm-zero in the minimization). In the adaptive gradient-based method, the missing samples/measurements are considered as the minimization variables. The available samples values are known and fixed. Obviously, the number

of variables in the minimization process is equal to the number of missing samples/measurements in the observation domain. This approach is possible when the common signal transforms are the domains of signal sparsity. Then missing and available samples/measurements form a complete set of samples/measurements.

The discrete Fourier transform (DFT), as the most important signal transform, is considered in this paper as the sparsity domain of the signal. A theorem for the uniqueness of the reconstructed solution, based on the missing sample variations, is presented. Two forms of the theorem are presented. One stating the uniqueness condition for a given missing sample transformation matrix and the other providing the uniqueness check if a sparse signal is already recovered using a reconstruction algorithm. The solution is unique in the sense that the variation of the missing sample values can not produce another signal of the same or lower sparsity. The theorems provide an easy and computationally efficient uniqueness check.

The paper is organized as follows. After the introduction, the uniqueness theorems and corollaries are defined and illustrated on examples. The proofs are presented in Section 3. The worst case signal is derived and related to the group delay function in Section 4. Theoretical results are demonstrated on simple illustrative and statistical examples as well.

II. ON THE RECONSTRUCTED SIGNAL UNIQUENESS

Consider a signal $x(n)$ with $n \in \mathbb{N} = \{0, 1, 2, \dots, N - 1\}$. Assume that Q of its samples at the positions $q_m \in \mathbb{N}_Q = \{q_1, q_2, \dots, q_Q\}$ are missing/omitted. The signal is sparse in the DFT domain, with sparsity s . The reconstruction goal is to get $x(n)$, for all $n \in \mathbb{N}$ using available samples at $n \in \mathbb{M} = \mathbb{N} \setminus \mathbb{N}_Q$. We will consider a new signal of the form

$$x_a(n) = x(n) + z(n)$$

where $z(n) \equiv 0$ for the available signal positions $n \in \mathbb{M}$ and $z(n)$ may take arbitrary values at the positions of missing samples $n = q_m \in \mathbb{N}_Q = \{q_1, q_2, \dots, q_Q\}$. The DFT of this signal is

$$X_a(k) = X(k) + Z(k) \\ = \sum_{i=1}^s \sigma_i \delta(k - k_{0i}) + \sum_{m=1}^Q z(q_m) e^{-j \frac{2\pi}{N} q_m k / N}.$$

University of Montenegro, Podgorica, ljubisa@ac.me, milos@ac.me

This is preprint version of article published in Mathematical Problems in Engineering (Hindawi Publishing Corporation), Article ID 629759, <http://www.hindawi.com/journals/mpe/aip/629759/>

Positions of nonzero values in $X(k)$ are $k_{0i} \in \mathbb{K}_s = \{k_{01}, k_{02}, \dots, k_{0s}\}$ with $X(k_{0i}) = \sigma_i$. In the minimization process the values of missing samples of $x_a(n) = x(n) + z(n)$ for $n \in \mathbb{N}_Q$ are considered as variables. The goal of the reconstruction process is to get $x_a(n) = x(n)$, or $z(n) = 0$ for all $n \in \mathbb{N}$. This goal should be achieved by minimizing the sparsity of the signal transform $X_a(k)$. Existence of the unique solution of this problem depends on the number of missing samples, their positions, and the signal itself.

First, assume that the signal can take any form, including the worst possible one. Then the number of missing samples and their positions will be considered only. The uniqueness, in this case, means that if a signal with the transform $X(k)$ of sparsity s , is obtained using a reconstruction method, with a given set of missing samples, then there is no other signal of the same or lower sparsity that satisfy the given set of available samples values, using the same set of missing samples as variables.

Theorem 1 Consider a signal $x(n)$ that is sparse in the DFT domain with unknown sparsity. Assume that the signal length is $N = 2^r$ samples and that Q samples are missing at the instants $q_m \in \mathbb{N}_Q$. Assume that the reconstruction is performed and that the DFT of reconstructed signal is of sparsity s . The reconstruction result is unique if the inequality

$$s < N - \max_{h=0,1,\dots,r-1} \{2^h (Q_{2^h} - 1)\} - s$$

holds. Integers Q_{2^h} are calculated as

$$Q_{2^h} = \max_{b=0,1,\dots,2^h-1} \{\text{card}\{q : q \in \mathbb{N}_Q \text{ and } \text{mod}(q, 2^h) = b\}\} \quad (1)$$

Example: Consider a signal with $N = 2^5 = 32$ and $Q = 9$ missing samples at

$$q_m \in \mathbb{N}_Q = \{2, 3, 8, 13, 19, 22, 23, 28, 30\}.$$

Using the theorem, we will find the sparsity limit s when we are able to claim that the reconstructed sparse signal is unique for any signal form.

-For $h = 0$ we use $Q_{2^0} = Q$ and get $2^0 (Q_{2^0} - 1) - 1 = (Q - 1) - 1 = 9$.

-For $h = 1$, the number Q_{2^1} is the greater value of

$$\begin{aligned} \text{card}\{q : q \in \mathbb{N}_Q \text{ and } \text{mod}(q, 2) = 0\} = \\ \text{card}\{2, 8, 22, 28, 30\} = 5 \end{aligned}$$

$$\begin{aligned} \text{card}\{q : q \in \mathbb{N}_Q \text{ and } \text{mod}(q, 2) = 1\} = \\ \text{card}\{3, 13, 19, 23\} = 4, \end{aligned}$$

i.e., the maximal number of even or odd positions of missing samples. Thus $Q_{2^1} = \max\{5, 4\} = 5$ with $2^1 (Q_{2^1} - 1) = 8$.

-Next Q_{2^2} is calculated as the maximal number of missing samples whose distance is multiple of 4. For various initial counting positions $b = 0, 1, 2, 3$ the numbers of missing samples with distance being multiple of 4 are 2, 1, 3, and 3, respectively. Then $Q_{2^2} = \max\{2, 1, 3, 3\} = 3$ with $2^2 (Q_{2^2} - 1) = 8$.

-For Q_{2^3} the number of missing samples at distances being multiple of 8 are found for various $b = 0, 1, 2, 3, 4, 5, 6, 7$. The

value of Q_{2^3} is 2 with $2^3 (Q_{2^3} - 1) = 8$.

-Finally we have two samples at distance 16 (samples at the positions $q_2 = 3$ and $q_5 = q_2 + N/2$) producing $Q_{2^4} = Q_{16} = 2$ with $2^4 (2 - 1) = 16$.

The reconstructed signal of sparsity s is unique if

$$\begin{aligned} s < N - \max_{h=0,1,2,3,4} \{2^h (Q_{2^h} - 1)\} - s \\ s < 32 - \max\{9, 8, 8, 8, 16\} - s \\ s < 32 - 16 - s \end{aligned}$$

or

$$s < 8.$$

The theorem considers general signal form. It includes the case when the amplitudes of signal components are related to each other and related to the missing sample positions. The specific signal form required by the theorem, to reach its bound, is analyzed in Section IV. Since this kind of relation is a zero-probability event, we will define a corollary, neglecting the probability that the signal values are dependent to each other and related to missing sample positions at the same time.

Corollary 2 Consider the signal $x(n)$ that is sparse in the DFT domain. Assume that signal length is $N = 2^r$ samples and that Q samples are missing at the instants $q_m \in \mathbb{N}_Q$. Also assume that the reconstruction is performed and that the DFT of reconstructed signal is of sparsity s . Assume that the amplitudes of signal components are arbitrary with arbitrary phases so that the case when all of them can be related to the values defined by using the missing sample positions is a zero-probability event. The reconstruction result is not unique if the inequality

$$s \geq N - \max_{h=0,1,\dots,r-1} \{2^h (Q_{2^h} - 1)\} - 1$$

holds. Integers Q_{2^h} are calculated in the same way as in the Theorem 1.

Example: Consider a signal with $N = 2^5 = 32$ and $Q = 9$ missing samples at

$$q_m \in \mathbb{N}_Q = \{2, 3, 8, 13, 19, 22, 23, 28, 30\}.$$

The sparsity limit s when we are able to claim that the reconstructed sparse signal is not unique is

$$\begin{aligned} s \geq N - \max_{h=0,1,2,3,4} \{2^h (Q_{2^h} - 1)\} - 1 \\ s \geq 32 - \max\{9, 8, 8, 8, 16\} - 1 \\ s \geq 15. \end{aligned}$$

Pseudo-code for uniqueness check according to Theorem 1 and Corollary 2 is presented in Algorithm 1. Note that in Corollary 2 we used the condition that the reconstruction result is nonunique (instead of the condition that the reconstruction result is unique in Theorem 1) since the zero-probability events are included here.

Corollary 2 provides the uniqueness test for the given positions of unavailable samples. In the cases with $h > 0$ it exploits the periodic structure of the transformation matrix of missing samples. The periodical form assumes that the

Algorithm 1 Sparsity limits – Theorem 1 and Corollary 2

Require:

- Set of missing sample positions \mathbb{N}_Q
- Total number of signal samples N , $N = 2^r$

```

1:  $K_T \leftarrow N$ 
2:  $K_C \leftarrow N$ 
3: for  $h \leftarrow 1, r - 1$  do
4:   for  $b \leftarrow 0, 2^h - 1$  do
5:      $Q_b \leftarrow \text{card}\{q : q \in \mathbb{N}_Q \text{ and } \text{mod}(q, 2^h) = b\}$ 
6:   end for
7:    $Q_{2^h} \leftarrow \max_b Q_b$ 
8:   if  $K_T > (N - 2^h(Q_{2^h} - 1))/2$  then
9:      $K_T \leftarrow (N - 2^h(Q_{2^h} - 1))/2$ 
10:   end if
11:   if  $K_C > N - 2^h(Q_{2^h} - 1) - 1$  then
12:      $K_C \leftarrow N - 2^h(Q_{2^h} - 1) - 1$ 
13:   end if
14: end for
    
```

Output: K_T, K_C

- Every solution with sparsity $s < K_T$ is unique.
- Solution with sparsity $s < K_C$ is unique with probability one, excluding zero-probability event (when the amplitudes of signal components are related to each other with a relation defined by the missing sample positions).

positions of possible zero values in $Z(k)$ do not interfere with the signal nonzero value positions. This is possible in the worst case analysis. For example, with two missing samples at the positions $z(q_m)$ and $z(q_m + N/2)$, the reconstruction process assumes that there are $N/2 - 1$ zero values in $Z(k)$ and that the same number of zero values is preserved in $X(k) + Z(k)$. This event can occur if we assume that all nonzero values of $X(k)$ have the same structure as $Z(k)$. In this specific case, it means that all of the signal nonzero coefficients are either on odd or even positions in the frequency domain.

Numerical example: Signal with $N = 64$ samples is considered. The signal sparsity is varied from $s = 1$ to $s = N - 1$. For each signal sparsity s number of missing samples is varied from $Q = 1$ to $Q = N - 1$. For each pair (s, Q) 1000 trials are performed with Q randomly positioned missing samples. Uniqueness is checked by Theorem 1 and Corollary 2. Percentage of trials where uniqueness is guaranteed is presented in Fig. 1. We can clearly see two regions one where uniqueness was achieved in each trial ($P = 100\%$) and the other where the solution was always nonunique ($P = 0\%$). In the transition between these two regions the uniqueness highly depends on the missing sample positions, producing $0\% < P < 100\%$.

We see that there is a sharp transition, for example at $Q = 16$ (for $N - Q = 48$), from $P(s, Q) = P(15, 16) \cong 1$ when $s = 15$ to $P(s, Q) = P(16, 16) \cong 0$ for $s = 16$

(marked with white dots in Fig. 1(left)). It means that the difference of probabilities $P(15, 16) - P(16, 16)$ is almost 1. Let us explain this effect. Consider Theorem 1 condition $s < N - 2^h(Q_{2^h} - 1) - s$ for $h = 0, 1, \dots, r - 1$. If $32 = 64 - 2^h(Q_{2^h} - 1)$ the solution will be unique for $s = 15$ and nonunique for $s = 16$. This equality is satisfied for $Q_{32} = 2$, $Q_{16} = 3$, $Q_8 = 5$, and $Q_4 = 9$. It means that at least one of the following holds: 1) There are $Q_{32} = 2$ samples at distance 32, 2) There are $Q_{16} = 3$ samples at the distance being multiple of 16, 3) There are $Q_8 = 5$ samples at the distance being multiple of 8, 4) There are $Q_4 = 9$ samples at the distance being multiple of 4. Probability that among $Q = 16$ samples out of $N = 64$ there are 2 samples at the distance 32 is $P = 0.92$. Since other events are with lower probabilities, this one is sufficient to explain the sharp change in $P(s, Q)$. We may write $P(15, 16) - P(16, 16) \sim 0.92$.

After signal reconstruction, we are in a position to additionally specify the uniqueness requirements using the reconstructed signal. When a sparse signal is reconstructed we want to check the uniqueness of this signal. It means that the signal $x(n)$, with its transform $X(k)$ which is nonzero at $k_i \in \mathbb{K}_s = \{k_{01}, k_{02}, \dots, k_{0s}\}$, is obtained and we want to check if there is another signal $X(k) + Z(k)$ with the same or lower sparsity, where $Z(k)$ is the DFT of arbitrary values of samples at the missing sample positions. The positions of nonzero values in $X(k)$ are not arbitrary, while the positions of zero and nonzero values in $Z(k)$ could change to produce minimal possible sparsity of $X(k) + Z(k)$. In the previous example with two missing samples $z(q_m)$ and $z(q_m + N/2)$, when the signal is already recovered, it means that we can not assume that all $\{k_{01}, k_{02}, \dots, k_{0s}\}$ are either odd or even. They are given since we have already reconstructed a sparse signal.

Theorem 3 Consider the signal $x(n)$ that is sparse in the DFT domain with unknown sparsity. Assume that the signal length is $N = 2^r$ samples and that Q samples are missing at the instants $q_m \in \mathbb{N}_Q$. Also assume that the reconstruction is performed and that the DFT of reconstructed signal is of sparsity s . Assume that the positions of the reconstructed nonzero values in the DFT are $k_{0i} \in \mathbb{K}_s = \{k_{01}, k_{02}, \dots, k_{0s}\}$ Reconstruction result is unique if the inequality

$$s < N - \max_{h=0,1,\dots,r-1} \{2^h(Q_{2^h} - 1) - s + 2S_{2^{r-h}}\}$$

holds. Integers Q_{2^h} and $S_{2^{r-h}}$ are calculated as

$$Q_{2^h} = \max_{b=0,1,\dots,2^h-1} \{\text{card}\{q : q \in \mathbb{N}_Q, \text{mod}(q, 2^h) = b\}\}$$

$$S_{2^{r-h}} = \sum_{l=1}^{Q_{2^h}-1} P_h(l)$$

$$P_h(l) = \text{sort}_{b=0,1,\dots,2^{r-h}-1} \{\text{card}\{k : k \in \mathbb{K}_s, \text{mod}(k, 2^{r-h}) = b\}\}$$

where $P_h(1) \leq P_h(2) \leq \dots \leq P_h(2^{r-h})$.

Note that for $S_{2^{r-h}} = 0$ this Theorem reduces to the Theorem 1. For the DFT values equally distributed over all positions, this Theorem produces result close to $s \geq N - Q$.

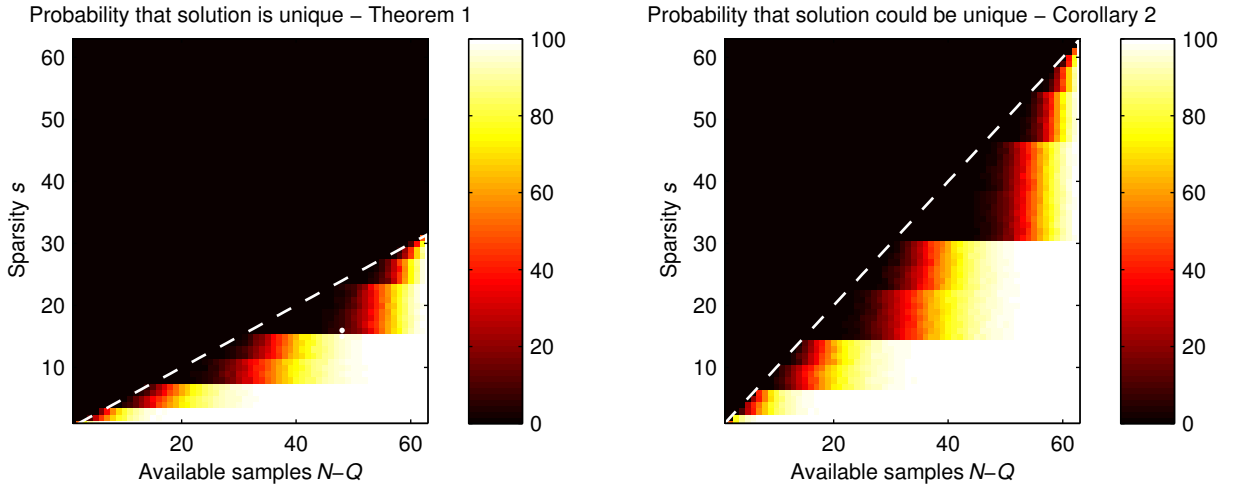


Fig. 1. Uniqueness probability calculated for various sparsity s and number of available samples $(N - Q)$ by using Theorem 1 (left) and Corollary 2 (right) in 1000 random realizations. Lines $s = N - Q$ and $s = (N - Q)/2$ are presented as guidelines. In this case the uniqueness is checked without using information about nonzero DFT coefficient positions in the reconstructed signal.

Corollary 4 Consider the signal $x(n)$ that is sparse in the DFT domain with unknown sparsity. Assume that signal length is $N = 2^r$ samples and that Q samples are missing at the instants $q_m \in \mathbb{N}_Q$. Also assume that the reconstruction is performed and that the DFT of reconstructed signal is of sparsity s . Assume that the positions of the reconstructed nonzero values in the DFT are $k_{0i} \in \mathbb{K}_s = \{k_{01}, k_{02}, \dots, k_{0s}\}$. Assume that the amplitudes of signal components are arbitrary with arbitrary phases so that the case when all of them can be related to the values defined by using the missing sample positions is a zero-probability event. Reconstruction result is not unique if the inequality

$$s \geq N - \max_{h=0,1,\dots,r-1} \{2^h (Q_{2^h} - 1) - 1 + S_{2^{r-h}}\}$$

holds. Integers Q_{2^h} and $S_{2^{r-h}}$ are calculated as in the Theorem 3. The case when all of signal components can be related to the values defined by using the missing sample positions is considered here.

Pseudo-code for uniqueness check according to Theorem 3 and Corollary 4 is presented in Algorithm 2.

Example: Consider a signal with $N = 32$ and $Q = 9$ missing samples at

$$q_m \in \mathbb{N}_Q = \{2, 3, 8, 13, 19, 22, 23, 28, 30\}.$$

Assume that with these missing samples we have reconstructed signals with nonzero DFT values at the positions

- a) $\mathbb{K}_s = \{1, 3, 5, 7, 9, 11, 13, 15, 17, 21, 23, 25, 27, 29, 31\}$,
- b) $\mathbb{K}_s = \{1, 3, 5, 9, 13, 17, 21, 29, 31, 2, 4, 8, 12, 16, 20, 24, 30\}$.

By testing these two signals, we get the following decisions. According to Theorem 1 we can not claim uniqueness in either of these cases since $s = 15$ in the first case and $s = 17$ in the second case. Both are greater than the Theorem 1 bound $s < 8$. The same holds for Corollary 2 since both are $s \geq 15$. By testing these results with Theorem 3 we

get that that in case a) the solution is nonunique. It is due to a very specific form of the reconstructed signal with all components being found at the odd frequency positions. Since the sparsity was defined by periodicity 16 in $q_m \in \mathbb{N}_Q$, then variations of two signal samples $z(q_2 = 3)$ and $z(q_5 = 19)$ can produce a signal $X(k) + Z(k)$ with the same sparsity as the reconstructed signal. These two samples, as variables, are able to produce many $(N/2)$ zero values in $Z(k)$ either at odd or even positions in frequency (Section 4). In this case, they are at even positions of $X(k) + Z(k)$. However, in signal b) that is not the case. Nonzero values are distributed over both even and odd frequency positions. Although sparsity of this signal is $s = 17$ the reconstruction is unique. The distribution of nonzero values in the reconstructed $X(k)$ is such that by varying two samples $z(q_2 = 3)$ and $z(q_5 = 19)$ we can not produce a signal $X(k) + Z(k)$ of the same or lower sparsity with nonzero $z(q_2 = 3)$ and $z(q_5 = 19)$. The limit, in this case, is defined by the lower periodicity in $z(q)$ than $N/2$. Thus, if we obtain this signal using a reconstruction algorithm the solution is unique.

Example: Consider a signal with $N = 1024$ and $Q = 512$ missing samples at $q_m \in \mathbb{N}_Q = \{0, 2, 4, \dots, 1022\}$. The reconstructed signal is at the frequencies: a) $\mathbb{K}_s = \{3\}$, b) $\mathbb{K}_s = \{3, 515\}$. We can easily check that in all cases with Theorem 1, Corollary 2 and Theorem 3, the reconstruction is nonunique although $s = 1$ or $s = 2$ is much smaller than the available number of samples $N - Q = 512$. The answer is obtained almost immediately, since the computational complexity of Theorem 1, Corollary 2 and Theorem 3, is of order $O(N)$.

Numerical example: Signal with $N = 64$ samples is considered. The signal sparsity is varied from $s = 1$ to $s = N - 1$. For each signal sparsity s the number of missing samples is varied from $Q = 1$ to $Q = N - 1$. For each pair (s, Q) 1000 trials are performed with randomly positioned s nonzero DFT values and randomly positioned Q missing samples. In each

Algorithm 2 Uniqueness check – Theorem 3 and Corollary 4

Require:

- Set of missing sample positions \mathbb{N}_Q
- Set of nonzero values in the reconstructed DFT \mathbb{K}_s
- Total number of signal samples N , $N = 2^r$

```

1:  $R_1 \leftarrow 1$ 
2:  $R_2 \leftarrow 1$ 
3:  $K \leftarrow \text{card } \mathbb{K}_s$ 
4: for  $h \leftarrow 1, r - 1$  do
5:   for  $b \leftarrow 0, 2^h - 1$  do
6:      $Q_b \leftarrow \text{card}\{q : q \in \mathbb{N}_Q \text{ and } \text{mod}(q, 2^h) = b\}$ 
7:   end for
8:    $Q_{2^h} \leftarrow \max_b Q_b$ 
9:   for  $b \leftarrow 0, 2^{r-h} - 1$  do
10:     $P_b \leftarrow \text{card}\{k : k \in \mathbb{K}_s \text{ and } \text{mod}(k, 2^{r-h}) = b\}$ 
11:   end for
12:   Sort array  $P$  in non-decreasing order
13:    $S_{2^{r-h}} \leftarrow \sum_{l=1}^{Q_{2^h}-1} P_{\text{sorted}}(l)$ 
14:   if  $K \geq N - 2^h(Q_{2^h} - 1) - K + 2S_{2^{r-h}}$  then
15:      $R_2 \leftarrow 0$ 
16:   end if
17:   if  $K \geq N - 2^h(Q_{2^h} - 1) - 1 + S_{2^{r-h}}$  then
18:      $R_1 \leftarrow 0$ 
19:   end if
20: end for
21:  $R \leftarrow R_1 + R_2$ 
    
```

Output: R

- $R = 2$ when the considered solution is unique.
- $R = 1$ when the considered solution is unique with probability one excluding zero-probability event (when amplitudes of the signal components are related to each other with a relation defined by missing sample positions).
- $R = 0$ when the considered solution is not unique.

trial the uniqueness is checked by Theorem 3 and Corollary 4. Percentage of trials where the uniqueness is guaranteed is presented in Fig. 2. We can clearly see two regions one where the uniqueness was achieved in each trial ($P = 100\%$) and the other region where the solution was nonunique in each trial ($P = 0\%$). The transition between regions is quite sharp. In this transition region, for given (s, Q) , the uniqueness highly depends on the positions of DFT values and missing samples producing $0\% < P < 100\%$. The transition region is wider for the Theorem 3 comparing to transition region for Corollary 4.

In Fig. 3, regions where the probability is higher than 99% are presented. The first region is defined by Theorem 3, which guaranties uniqueness for any signal. The second region is defined by Corollary 4 and it guaranties uniqueness with a high probability (this region is signal dependent). These regions are

combined in the third subplot.

III. PROOFS OF THE THEOREMS AND COROLLARIES

A. Proof of Theorem 1

Consider the DFT of a signal $x(n)$ with N samples. For the presentation simplicity we assumed common $N = 2^r$ although the results can be generalized for any N . Assume that M available samples are at the instants $n_i \in \mathbb{M} = \{n_1, n_2, \dots, n_M\}$ and the missing samples are at the instants $q_m \in \mathbb{N}_Q = \{q_1, q_2, \dots, q_Q\}$ and $Q = N - M$. Assume that the reconstruction process is done and the obtained signal is sparse in the DFT domain. Its sparsity is s with nonzero coefficients $k_{0i} \in \mathbb{K}_s = \{k_{01}, k_{02}, \dots, k_{0s}\} \subset \{0, 1, 2, 3, \dots, N - 1\}$. Under this assumption the signal $x(n)$ is

$$x(n) = \sum_{i=1}^s \frac{1}{N} \sigma_i e^{j2\pi n k_{0i}/N}.$$

Amplitudes $|\sigma_i/N|$ of the signal components are arbitrary with arbitrary phases $\arg\{\sigma_i\}$.

Let us form a signal

$$x_a(n) = x(n) + z(n)$$

where $z(n) = 0$ for $n \in \mathbb{M}$ and takes arbitrary values $z(q_m)$ at the positions of missing samples $q_m \in \mathbb{N}_Q$. The DFT of this signal is

$$X_a(k) = \sum_{i=1}^s \sigma_i \delta(k - k_{0i}) + \sum_{m=1}^Q z(q_m) e^{-j2\pi q_m k/N}.$$

Denote the number of non-zero values in $X_a(k)$ as $\|\mathbf{X}_a\|_0$. The DFT of signal $z(n)$ is

$$Z(k) = \text{DFT}\{z(n)\} = \sum_{m=1}^Q z(q_m) e^{-j2\pi q_m k/N}$$

Denote number of nonzero values in $Z(k)$ as $N_Z = \|\mathbf{Z}\|_0$. The aim of minimization of $\|\mathbf{X}_a\|_0$ is to produce the smallest possible value $\|\mathbf{X}_a\|_0 = s$. If this is possible in trivial case $z(n) = 0$ only, then our solution is unique. If there exists a non-trivial solution for $z(n)$ with $\|\mathbf{X}_a\|_0 = s$ then we have two different solutions $x(n)$ and $x(n) + z(n)$ with the same sparsity s meaning that our solution is not unique.

The DFT of signal $z(n)$ can be written in a matrix form as

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-j2\pi q_1/N} & e^{-j2\pi q_2/N} & \dots & e^{-j2\pi q_Q/N} \\ \dots & \dots & \dots & \dots \\ e^{-j2\pi q_1 \frac{N-1}{N}} & e^{-j2\pi q_2 \frac{N-1}{N}} & \dots & e^{-j2\pi q_Q \frac{N-1}{N}} \end{bmatrix} \times \begin{bmatrix} z(q_1) \\ z(q_2) \\ \dots \\ z(q_Q) \end{bmatrix} = \begin{bmatrix} Z(0) \\ Z(1) \\ \dots \\ Z(N-1) \end{bmatrix} \quad (2)$$

Since we look for zeros in $Z(k)$, without loss of generality, we can rewrite the system (2) in the form normalized with the

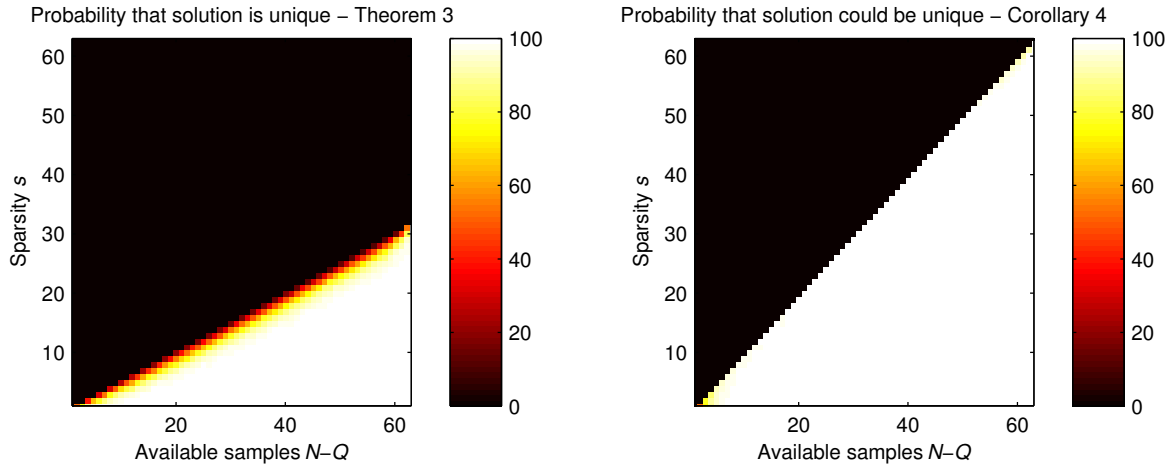


Fig. 2. Uniqueness probability calculated for various sparsity s and number of available samples $(N - Q)$ by using Theorem 3 (left) and Corollary 4 (right) in 1000 random realizations. In this case the uniqueness is checked after reconstruction using information about nonzero DFT coefficient positions in the reconstructed signal.

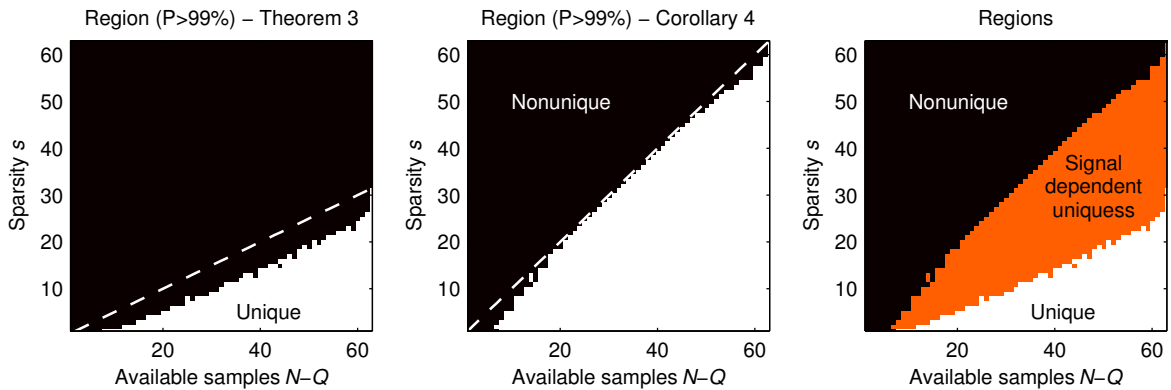


Fig. 3. Regions with uniqueness probability greater than 99% for various sparsity s and number of available samples $(N - Q)$ obtained with Theorem 3 (left), Corollary 4 (middle) and combined regions (right) in 1000 random realizations. Lines $s = N - Q$ and $s = (N - Q)/2$ are presented as guidelines.

first column corresponding to q_1 as

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & e^{-j2\pi q_{12}/N} & \dots & e^{-j2\pi q_{1Q}/N} \\ \dots & \dots & \dots & \dots \\ 1 & e^{-j2\pi q_{12} \frac{N-1}{N}} & \dots & e^{-j2\pi q_{1Q} \frac{N-1}{N}} \end{bmatrix} \times \begin{bmatrix} z(q_1) \\ z(q_2) \\ \dots \\ z(q_Q) \end{bmatrix} = \begin{bmatrix} Z(0) \\ Z(1)e^{j2\pi q_1/N} \\ \dots \\ Z(N-1)e^{j2\pi q_1(N-1)/N} \end{bmatrix} \quad (3)$$

where

$$q_{1n} = q_n - q_1.$$

In general, during the minimization, we have Q variables $z(q_m)$ (Q degrees of freedom). First assume that there is no common period (smaller than N) in columns of the transform matrix for the missing sample positions q_i . This case corresponds to pairwise coprime numbers $q_{12}, q_{13}, \dots, q_{1Q}$. Then Q variables $z(q_m)$ can be used to produce $Q - 1$ zeros in $Z(k)$ at the positions where there is no signal and, in the

worst possible case, to cancel out all s nonzero coefficients in $X(k)$. Therefore the largest possible number of zero values in $X(k) + Z(k)$ is $(Q - 1) + s$. The lowest possible number of nonzero values in $X(k) + Z(k)$ is $N - (Q - 1) - s$. If

$$N - (Q - 1) - s > s$$

then the considered solution is unique since only the trivial solution $z(n) = 0$ results in sparsity s and every nontrivial solution results in $\|\mathbf{X}_a\|_0 > s$. If we obtain the reconstructed signal with sparsity

$$s \geq \frac{N - Q + 1}{2}$$

then the solution is not unique. However, if $2s < N - Q + 1$ then it still does not mean that the solution is unique since we assumed that there is no periodicity in transform matrix (3). Next we will include all possible cases when some of the columns of transform matrix in (3) may have a common period lower than N .

The matrix of coefficients $Z(k)$ can not be periodic with

period smaller than Q . Thus, regarding to the period of the whole matrix the worst case would be the transformation matrix which repeats after exactly Q rows, where Q is divisor of N . Denote the number of repetitions as $R = N/Q$.

For periodic structure of transformation matrix in (3), with such a period that the rows are repeated R times within N , the largest number of zero values in $Z(k)$ is now $R(Q - 1)$. In addition, the nonzero values in $Z(k)$, in the worst possible case, can cancel all s nonzero DFT values of $X(k)$. Thus, in this case, the lowest possible number of nonzero values in $X(k) + Z(k)$ is $N - R(Q - 1) - s$. For the unique solution it should be greater than the signal sparsity, i.e., $N - R(Q - 1) - s > s$ or $N/(2Q) > s$. This is the result of the uncertainty principle in the DFT, presented in Subsection III-B.

The process does not end here. In the minimization process, we must also consider all subsets of missing samples $q_m \in \mathbb{N}_Q = \{q_1, q_2, \dots, q_Q\}$. Namely, it can happen that the worst case regarding the maximal number of zero values in $Z(k)$ is not the case with the full set of variables $z(q)$. It could happen that some subsets of the variables $q_m \in \mathbb{N}_Q$ may produce a higher number of zero values in $Z(k)$ than the whole set of variables. It means that the reconstruction algorithm may find some variables $z(q_i)$ to be zero valued and vary only a subset of remaining variables $z(q_i)$.

Subsets of Missing Samples: We have concluded that the periodicity R reduces the sparsity of $X(k) + Z(k)$, by increasing the number of possible zero values in $Z(k)$. Consider a general set of missing sample positions $q_m \in \mathbb{N}_Q = \{q_1, q_2, \dots, q_Q\}$. Then the algorithm for uniqueness should check periods for all possible subsets of missing samples. Assuming as in Theorem 1, without loss of generality, that $N = 2^r$ the cases are as follows:

1) Using all missing samples, assuming that there is no common period for all of them (in the sense of (3)), the unique solution is obtained if $s < N - (Q - 1) - s$. The cases with periodic matrix structure will be included in the steps that follows.

2) Consider minimal possible number of periods $R > 1$ for $N = 2^r$. It is $R = 2$ repetitions in the matrix of coefficients (3) with period $N/2$. Any subset of $\mathbb{N}_Q = \{q_1, q_2, \dots, q_Q\}$ containing only even or only odd positions can be considered as a set of minimization variables with period $N/2$. The number of missing samples $q_m \in \mathbb{N}_Q$ at even positions can be written as

$$N_E = \text{card}\{q : q \in \mathbb{N}_Q \text{ and } \text{mod}(q, 2) = 0\}.$$

The same should be done for odd positions in $q_m \in \mathbb{N}_Q$

$$N_O = \text{card}\{q : q \in \mathbb{N}_Q \text{ and } \text{mod}(q, 2) = 1\}.$$

Since we look for the worst case in our reconstruction algorithm, we choose the set with more variables (more degrees of freedom). It is

$$Q_2 = \max_{b=0,1} \{\text{card}\{q : q \in \mathbb{N}_Q \text{ and } \text{mod}(q, 2) = b\}\}.$$

Since the number of periods in matrix of coefficients is 2 for

these two subsets of missing samples it means that at most

$$2(Q_2 - 1)$$

zero values can be produced in $Z(k)$. It can be larger than $(Q - 1)$. Therefore beside the condition $s < N - (Q - 1) - s$ considered in 1) we have also to check

$$s < N - 2(Q_2 - 1) - s.$$

and use the worst case as the limit for s . It means that, at this point, we should compare Q and $2(Q_2 - 1)$.

Note that the largest possible Q_2 is $Q_2 = N/2$. In this case, the solution is unique if

$$2s < N - N + 1 = 1.$$

This corresponds to the case when all even (or all odd) signal samples are missing. Then we can not uniquely reconstruct a signal even for sparsity $s = 1$.

3) This analysis should be continued for all possible periods in N . For $N = 2^r$ the next possible period is $N/4$ with number of periods $R = 4$. Coefficients in the matrix (3) are periodic with $N/4$ if the distance between missing samples is four. Thus, we should divide set of all missing sample positions \mathbb{N}_Q into subsets where distances between q_i are multiples of 4, i.e., when $\text{mod}(q_i, 4)$ is a constant. There are 4 such subsets obtained for $\text{mod}(q_i, 4) = b$ where $b = 0, 1, 2, 3$. In the same way as in step 2), denote the cardinality of the largest subset by Q_4

$$Q_4 = \max_{b=0,1,2,3} \{\text{card}\{q : q \in \mathbb{N}_Q \text{ and } \text{mod}(q, 4) = b\}\}.$$

If we find Q_4 such that

$$4(Q_4 - 1) > 2(Q_2 - 1) \text{ and } 4(Q_4 - 1) > (Q - 1)$$

then it means that in $z(n)$ we can consider as variables (non-zero values) only the samples from the set containing q_i such that $\text{mod}(q_i, 4) = b$ producing Q_4 . Then the unique solution is obtained only if

$$s < N - 4(Q_4 - 1) - s$$

The worst case is when the positions of all missing samples are such that $Q_4 = N/4$. Then the solution is unique if $s < 2$. In the worst case with as many as $3N/4$ of available samples ($Q = Q_4 = N/4$) we may reconstruct only signals with sparsity $s = 1$.

4) Next possible period for $N = 2^r$ is $N/8$. The periodicity with period 8 should be considered for any subset of $q_m \in \mathbb{N}_Q = \{q_1, q_2, \dots, q_Q\}$ calculating

$$Q_8 = \max_{b=0,1,2,3,4,5,6,7} \{\text{card}\{q : q \in \mathbb{N}_Q \text{ and } \text{mod}(q, 8) = b\}\}.$$

If

$$8(Q_8 - 1) > \max\{(Q - 1), 2(Q_2 - 1), 4(Q_4 - 1)\}$$

then the sparsity further reduces to

$$s < N - 8(Q_8 - 1) - s$$

The worst case when the positions of all missing samples are such that $Q_8 = N/8$ the solution is unique for $s < 4$. I this

case even with $7N/8$ of available samples we may reconstruct only the signals with sparsity $s = 1, 2, 3$.

5) The process should continue for all possible periods $2^h < N$, $h = 4, 5, \dots, r-1$. We should calculate

$$Q_{2^h} = \max_{b=0,1,\dots,2^h-1} \{\text{card}\{q : q \in \mathbb{N}_Q, \text{mod}(q, 2^h) = b\}\}.$$

If it is such that

$$2^h (Q_{2^h} - 1) > \max \{(Q-1), 2(Q_2-1), 4(Q_4-1), \dots, 2^{h-1}(Q_{2^{h-1}}-1)\}$$

then the sparsity constraint for uniqueness is

$$s < N - 2^h (Q_{2^h} - 1) - s.$$

6) Summing up all the cases we get the theorem result that the uniqueness condition is

$$s < N - \max_{h=0,1,\dots,r-1} \{2^h (Q_{2^h} - 1)\} - s$$

In the final stage if just two special samples

$$q_i \in \mathbb{N}_Q \text{ and } q_{i+N/2} \in \mathbb{N}_Q$$

are missing then $Q_{2^{r-1}} - 1 = 1$ with the solution being non-unique if

$$s \geq N - N/2 - s$$

Thus special positions of two missing samples reduce the maximal number of components that can uniquely be detected to $N/4 - 1$.

7) This kind of proof can be generalized for any signal length N with possible periods of matrix obtained as divisors of N .

B. Uncertainty Principle and Theorem 1 Bounds

The bounds for sparsity are compared with the results following from the uncertainty relation for the DFT

$$Q \cdot N_Q \geq N.$$

where the number of nonzero values in $z(q)$ is denoted by Q and the number of the nonzero values in its DFT $Z(k)$ is N_Q . In the worst case, for a given Q and for the worst possible distribution of positions $q_m \in \mathbb{N}_Q = \{q_1, q_2, \dots, q_Q\}$ we have

$$N_Q = \frac{N}{Q}.$$

Consider first the case when we can not exclude the possibility that the signal DFT assume values related to the missing sample positions and the worst possible distribution of positions $q_m \in \mathbb{N}_Q = \{q_1, q_2, \dots, q_Q\}$. The maximal number of nonzero values is N_Q . In the worst case s of nonzero signal DFT values can cancel s nonzero values in $Z(k)$ meaning that the minimal number of nonzero values, in the worst case, is $N_Q - s$. Since it should not be the solution of our minimization problem, the case with all $z(q)$, producing the signal and its sparsity should be lower, $N_Q - s > s$ or

$$s < \frac{N_Q}{2} = \frac{N}{2Q}.$$

For arbitrary signal whose values can not be described by the missing sample positions, and for the worst possible distribution of positions we get $N_Q - 1 > s$ or

$$s < \frac{N - Q}{Q} = \frac{N}{Q} - 1.$$

These two cases are obtained through the previous analysis as the special worst cases of the Theorem 1 and Corollary 2 as well.

C. Proof of Corollary 2

In Theorem 1 we assumed that $Z(k)$ takes maximal possible number of $Q - 1$ zero values and that at the same time the remaining nonzero values of $Z(k)$ cancel out all signal components $X(k)$. This assumption is very unlikely since after we assumed the maximal possible number of $Q - 1$ zeros in $Z(k)$, we have only one remaining degree of freedom. It means that we can cancel out one signal component with one remaining variable and we assume that all other components have specific values so that they will also be canceled out. Consider now the case when signal components are not adjusted to these fixed values of $Z(k)$. Then, in reality, we can expect to cancel out only one component in $X(k)$ with one variable. Repeating Theorem 1 proof with

$$s \geq N - 2^h (Q_{2^h} - 1) - 1$$

for nonuniqueness instead of $s < N - 2^h (Q_{2^h} - 1) - s$ for uniqueness produces the proof of Corollary 2.

D. Proof of Theorem 3

Here we will start with two missing samples q_{m1} and q_{m2} at the distance $|q_{m1} - q_{m2}| = N/2 = 2^{r-1} = 2^h$. The worst case is $Q_{2^h} - 1 = 1$ with the uniqueness condition

$$s < N - N/2 - 1 = N/2 - 1$$

These special positions of two missing samples reduce the maximal number of components that can uniquely be detected to $s = N/2 - 2$. However in this case the positions of nonzero DFT values $k_{0i} \in \mathbb{K}_s = \{k_{01}, k_{02}, \dots, k_{0s}\}$ should be very specific. All of them should be on either even or odd positions, producing

$$\begin{aligned} S_{2^{r-h}} = S_{2^1} &= \min_{b=0,1} \{\text{card}\{k : k \in \mathbb{K}_s, \text{mod}(k, 2) = b\}\} \\ &= \min\{0, 2\} = 0 \end{aligned}$$

If that is not the case then

$$\begin{aligned} S_{2^1} &= \min_{b=0,1} \{\text{card}\{k : k \in \mathbb{K}_s, \text{mod}(k, 2) = b\}\} \\ &= \min\{1, 1\} = 1 > 0 \end{aligned}$$

will reduce the count of zero values that can be achieved in the worst case in $X(k) + Z(k)$. Since S_2 of the signal nonzero coefficients are not at the positions of nonzero values of $Z(k)$ then, in the best possible case, only $(s - S_2)$ out of s nonzero values of the DFT can be canceled out by the predetermined missing sample values (as in Theorem 1). In addition to $N - 2^{r-1} (Q_{2^{r-1}} - 1) - (s - S_2)$ nonzero values of $X(k) + Z(k)$

at the nonzero positions of $Z(k)$, there are S_2 nonzero values of $X(k)$ positioned at the zero-value positions of $Z(k)$. They can not be canceled by $Z(k)$. The uniqueness condition for $h = r - 1$ will therefore require

$$\begin{aligned} s &< N - 2^{r-1} (Q_{2^{r-1}} - 1) - (s - S_{2^1}) + S_{2^1} \\ s &< N - 2^h (Q_{2^h} - 1) - (s - S_{2^{r-h}}) + S_{2^{r-h}}. \end{aligned}$$

This correction of the uniqueness condition with S_{2^1} should be done only if $Q_{2^h} = Q_{N/2} = Q_{2^{r-1}} = 2$. Denote with $P_h(l)$ the sorted array

$$P_h(l) = \text{sort}_{b=0,1} \{ \text{card}\{k : k \in \mathbb{K}_s \text{ and } \text{mod}(k, 2^{r-h}) = b\} \}.$$

The correction $S_{2^1} = S_{2^{r-h}}$ with $h = r - 1$ can be calculated as

$$S_{2^{r-h}} = \sum_{l=1}^{Q_{2^h}-1} P_h(l).$$

For $Q_{2^h} = Q_{2^{r-1}} = Q_{N/2} = 1$ (when there are no two samples at the distance $N/2$) the upper summation limit will be 0 and this kind of correction will not be done, $S_{2^1} = 0$. Note that the sum of all $P_h(l)$, $l = 1, 2, \dots, 2^{r-h}$ is equal to the signal sparsity s .

In the same way, for $h = r - 2$ (missing samples at a distance being multiple of $N/4$) the period of $Z(k)$ is 4. In the worst case when $Q_{N/4} = Q_{2^{r-2}} = 4$, the maximal number of nonzero values in $X(k)$ at a distance being multiple of 4 can be canceled out (with nonzero values of $Z(k)$). The remaining nonzero values in the DFT of signal will be on the zero positions of $Z(k)$ and can not be canceled out. Note that if $Q_{N/4} = Q_{2^{r-2}} = 3$ then two nonzero values in $Z(k)$ will remain (with 3 variables we can make only 2 zero values in each period of $Z(k)$). The total number of the nonzero values of $X(k)$ that can be canceled is the sum of two largest numbers in $P_h(l)$ defined by $P_h(4) + P_h(3) = s - P_h(1) - P_h(2)$, where

$$P_h(l) = \text{sort}_{b=0,1,2,3} \{ \text{card}\{k : k \in \mathbb{K}_s \text{ and } \text{mod}(k, 4) = b\} \}.$$

The number of remaining nonzero DFT values that can not be canceled out is

$$S_{2^2} = P_h(1) + P_h(2) = \sum_{l=1}^{Q_{2^{r-2}}-1} P_h(l).$$

Note that if $Q_{2^h} = Q_{N/4} = Q_{2^{r-2}} = 2$ then three nonzero values exists in one period of $Z(k)$ (only one zero value of $Z(k)$ can be obtained within a period) meaning that the total number of signal values that can be canceled out is $P_h(4) + P_h(3) + P_h(2) = s - P_h(1)$. The number of remaining nonzero values of $X(k)$ is $S_{2^2} = P_h(1) = \sum_{l=1}^{Q_{2^{r-2}}-1} P_h(l)$.

The uniqueness condition is that

$$s < N - 2^h (Q_{2^h} - 1) - (s - S_{2^{r-h}}) + S_{2^{r-h}} \quad (4)$$

with

$$\begin{aligned} P_h(l) &= \\ S_{2^{r-h}} &= \sum_{l=1}^{Q_{2^h}-1} P_h(l). \end{aligned} \quad (5)$$

and $2^{r-h} = 4$ for $h = r - 2$. The same analysis is done for all h , using (4)-(5). The value of s satisfying (4) for all h produces Theorem 3 statement.

E. Proof of Corollary 4

This Corollary follows from Theorem 3 neglecting the probability that several DFT coefficients can be canceled out by predetermined values of the missing samples. Then instead of $(s - S_{2^{r-h}})$ in (4) we have just one sample and

$$s \geq N - \max_{h=0,1,\dots,r-1} \{ 2^h (Q_{2^h} - 1) - 1 + S_{2^{r-h}} \}$$

is used for nonuniqueness instead of the uniqueness condition

$$s < N - \max_{h=0,1,\dots,r-1} \{ 2^h (Q_{2^h} - 1) - (s - S_{2^{r-h}}) + S_{2^{r-h}} \}.$$

IV. THE WORST CASE SIGNAL FORM

For the worst case, defined by the Theorem 1, the set of possible amplitudes and phases of signal components is related to the missing sample positions. For the missing sample positions $q_m \in \mathbb{N}_Q = \{q_1, q_2, \dots, q_Q\}$ used in the minimization process, the worst case in the minimization process is when the period of the transformation matrix is such that it repeats immediately after Q samples, in the sense described in the proof of Theorem 1. Then the minimal number of nonzero values in DFT $Z(k)$ is $N_Q = N/Q$. It means that Q variables $z(q_1), z(q_2), \dots, z(q_Q)$ can be determined such that $Q - 1$ values of $Z(k)$ within $k = 0, 1, 2, \dots, Q - 1$ are zero-valued and only one $Z(k)$ is nonzero. In the worst case, this scenario repeats immediately after Q values (it repeats N/Q times). The maximal number of zero values in $Z(k)$ is then $(Q-1)N/Q$. The number of nonzero values (sparsity) in $Z(k)$ is N/Q .

Now we will investigate the form of the signal DFT so that the Theorem 1 sparsity bound holds with equality sign. The worst case assumes existence of the maximal number of zeros in $Z(k)$ and one nonzero value of $Z(k)$ in each period N/Q . It also assumes that all s signal components can be canceled out by this nonzero value of $Z(k)$ and corresponding periodic nonzero values $Z(k + iQ)$. The maximal number of zeros in $Z(k)$ within one period defines values of all nonzero values of variables $z(n)$ in the time domain (with a possibility to cancel out one signal component in $X(k) + Z(k)$ within the considered period).

Let us consider subsets of Q equations defined by (2). The first subset will be written for frequencies $k = 0, 1, 2, \dots, Q-1$, the second for $k = Q, Q+1, \dots, 2Q-1$, and so on until the last one for $k = N-Q, N-Q+1, \dots, N-1$. Assume the signal DFT coefficient $X(k_{0i})$ (the one that we want to use along with $Q-1$ zero values of $Z(k)$ to calculate variables

$$\begin{bmatrix} e^{-j2\pi q_1 iQ/N} & e^{-j2\pi q_2 iQ/N} & \dots & e^{-j2\pi q_Q iQ/N} \\ e^{-j2\pi q_1 (iQ+1)/N} & e^{-j2\pi q_2 (iQ+1)/N} & \dots & e^{-j2\pi q_Q (iQ+1)/N} \\ \dots & \dots & \dots & \dots \\ e^{-j2\pi q_1 k_{0i}/N} & e^{-j2\pi q_2 k_{0i}/N} & \dots & e^{-j2\pi q_Q k_{0i}/N} \\ \dots & \dots & \dots & \dots \\ e^{-j2\pi q_1 (iQ+Q-1)/N} & e^{-j2\pi q_2 (iQ+Q-1)/N} & \dots & e^{-j2\pi q_Q (iQ+Q-1)/N} \end{bmatrix} \begin{bmatrix} z(q_1) \\ z(q_2) \\ \dots \\ z(q_Q) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ -X(k_{0i}) \\ \dots \\ 0 \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} z(q_1) \\ z(q_2) \\ \dots \\ z(q_Q) \end{bmatrix} = \begin{bmatrix} e^{-j2\pi q_1 iQ/N} & e^{-j2\pi q_2 iQ/N} & \dots & e^{-j2\pi q_Q iQ/N} \\ e^{-j2\pi q_1 (iQ+1)/N} & e^{-j2\pi q_2 (iQ+1)/N} & \dots & e^{-j2\pi q_Q (iQ+1)/N} \\ \dots & \dots & \dots & \dots \\ e^{-j2\pi q_1 k_{0i}/N} & e^{-j2\pi q_2 k_{0i}/N} & \dots & e^{-j2\pi q_Q k_{0i}/N} \\ \dots & \dots & \dots & \dots \\ e^{-j2\pi q_1 (iQ+Q-1)/N} & e^{-j2\pi q_2 (iQ+Q-1)/N} & \dots & e^{-j2\pi q_Q (iQ+Q-1)/N} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ \dots \\ -X(k_{0i}) \\ \dots \\ 0 \end{bmatrix} \quad (7)$$

$z(q_1), z(q_2), \dots, z(q_Q)$) is within the subset of equations for $k = iQ, iQ + 1, \dots, iQ + Q - 1$. Then the solution of system (6) will produce maximal sparsity for this frequency range, i.e., $Z(k) + X(k) = 0$ for all considered frequencies $k = iQ, iQ + 1, \dots, iQ + Q - 1$. Solution for the missing sample values is given by (7)

Values of $z(q)$, obtained from this system, do not change for other $s - 1$ nonzero signal DFT values $X(k_{0m})$, $k_{0m} \neq k_{0i}$. With the previous system of equations and its solution for missing samples $z(q)$ we have used all degrees of freedom of Q -dimensional variable $z(q)$. More zero values in the DFT of the resulting signal $X(k) + Z(k)$ (lower sparsity of this signal) can be obtained only if the remaining signal DFT values $X(k_{0m})$, $k_{0m} \neq k_{0i}$ are canceled out, by chance. In the worst case, assumed by the Theorem 1, all remaining $s - 1$ nonzero DFT coefficients $k_{0m} \neq k_{0i}$ should be canceled out, with the already determined missing sample values $z(q)$. Since the transformation matrix is periodic, this will happen only if all remaining signal DFT coefficients assume very specific positions $k_{0m} = k_{0i} + r_m Q$ and specific values

$$\begin{aligned} X(k_{0m}) e^{j2\pi q_1 k_{0m}/N} &= X(k_{0i}) e^{j2\pi q_1 k_{0i}/N} \\ X(k_{0m}) &= e^{-j2\pi (k_{0m} - k_{0i}) Q q_1 / N} X(k_{0i}) \\ X(k_{0m}) &= e^{-j2\pi r_m Q q_1 / N} X(k_{0i}), \end{aligned} \quad (8)$$

since the periodicity is established with respect to q_1 .

For a given set of missing sample positions $q_m \in \mathbb{N}_Q = \{q_1, q_2, \dots, q_Q\}$ probability that all components of a measured signal assume specific positions with specific amplitudes and phases (related to the missing sample positions) defined by (8) is a zero-probability event.

A. Group Delay and Missing Sample Positions

Consider a signal $y(n)$, periodic with period N/Q , and defined for $n = 0, 1, \dots, N/Q - 1$ as

$$y(n) = \delta(n - n_0) X(k_{0i})$$

where n_0 is the remainder after q_1 is divided with N/Q . Note that the worst case requires that all missing sample positions q_m are at the positions being multiple of N/Q . It means that

the remainder after division of q_m with N/Q is a constant denoted by n_0 . The DFT of the considered signal is

$$Y(k) = e^{-j2\pi n_0 k / (N/Q)} X(k_{0i}) \quad k = 0, 1, \dots, N/Q - 1$$

It is interesting to note that the signal DFT values $X(k_{0m})$, defined with (8), are obtained as a subset of s values from $\{Y(k) \mid k = 0, 1, \dots, N/Q - 1\}$, i.e.,

$$X(k_{0m}) \in \left\{ e^{-j2\pi n_0 k / (N/Q)} X(k_{0i}) \mid k = 0, 1, \dots, \frac{N}{Q} - 1 \right\}$$

In the worst case, the DFT values of signal should be s samples of a full DFT of a periodic signal which would have group delay coinciding with the missing sample positions $q_m \in \mathbb{N}_Q = \{q_1, q_2, \dots, q_Q\}$, since $q_m = q_n + l \frac{N}{Q}$ in this case. It means that if the missing samples produce a periodic structure, then the signal values should follow this structure as well.

Example: For the signal with $N = 32$ and missing samples at

$$q_m \in \mathbb{N}_Q = \{2, 3, 8, 13, 19, 22, 23, 28, 30\}$$

the limit for sparsity s (when we claim that the reconstructed sparse signal is unique, assuming that all signal amplitudes may be related to the missing sample positions) is $s < 8$. In this example, we will show what properties a signal must satisfy in the limit case $s = 8$ so that the solution is not unique. To simplify the notation assume that one DFT value of reconstructed signal is $X(5) = 2$.

The limit of sparsity $s = 8$ is obtained in the first example with $Q_{16} = 2$ and $16(2-1) = 16$. As explained, it corresponds to the missing sample positions $q_2 = 3$ and $q_5 = q_2 + N/2$. It means that the missing sample values of other samples $z(q_m)$ will be adjusted to their correct zero positions and only $z(q_2) = z(3) = z_3$ and $z(q_5) = z(19) = z_{19}$ will assume nonzero values. In this case the set of missing samples variables reduces to $q_m \in \mathbb{N}_Q = \{3, 19\}$ with $Q = Q_{16} = 2$. The DFT $Z(k)$ of such a signal is equal to

$$\begin{aligned} Z(k) &= z_3 e^{-j2\pi 3k/32} + z_{19} e^{-j2\pi 19k/32} \\ &= e^{-j2\pi 3k/32} (z_3 + (-1)^k z_{19}) \end{aligned}$$

$$\text{for } k = 0, 1, \dots, 31$$

In the worst case $Z(k)$ should have maximal possible number

of zeros. We conclude that either $z_3 = z_{19}$ or $z_3 = -z_{19}$ should hold, otherwise the sparsity of $Z(k)$ would be 32. In addition, $Z(k)$ should cancel out all signal components including assumed $X(5) = 2$. Since $z_3 = z_{19}$ would produce $Z(2k+1) = 0$ it would not be able to cancel $X(5)$. Therefore we must use $z_3 = -z_{19}$ with

$$Z(5) = e^{-j2\pi 15/32} (z_3 - z_{19}) = -X(5) = -2.$$

It means $z_3 = -z_{19} = -e^{j2\pi 15/32}$ and

$$Z(k) = \begin{cases} -2e^{-j2\pi(3k-15)/32} & \text{for odd } k \\ 0 & \text{for even } k. \end{cases}$$

In order to cancel all nonzero values of $X(k)$ they must be located at odd positions (where $Z(k)$ is nonzero)

$$\{5, k_{02}, k_{03}, k_{04}, k_{05}, k_{06}, k_{07}, k_{08}\}$$

and must be of opposite sign and equal amplitude to the corresponding (determined) values of $Z(k)$

$$X(k_{0i}) = -Z(k_{0i}) = 2e^{-j2\pi(3k_{0i}-15)/32} \quad \text{for } i = 2, 3, \dots, 8$$

resulting in

$$X(k) = \begin{cases} 2e^{-j2\pi(3k-15)/32} & \text{for } k \in \{5, k_{02}, k_{03}, k_{04}, k_{05}, \\ & k_{06}, k_{07}, k_{08}\} \\ 0 & \text{elsewhere.} \end{cases} \quad (9)$$

In this case sparsity of $X(k) + Z(k)$ is 8, the same as the sparsity of the $X(k)$. Two solutions of our minimization problem are signal $x(n) = \text{IDFT}[X(k)]$ and $x(n) + z(n)$ where

$$z(n) = \text{IDFT}[Z(k)] = \delta(n-3) - \delta(n-19).$$

Both of these signals have the same sparsity $s = 8$ and satisfy the same set of available samples.

However, if sampled signal $x(n)$ is not the signal of very specific form (9) then the solution of sparsity $s = 8$ will be unique for a given set of available samples. Then $z(n) = \delta(n-3) - \delta(n-19)$ will not be in position to cancel all 8 DFT values of signal and the sparsity of $X(k) + Z(k)$ will be 8 only for $z(n) = 0$, producing unique solution.

Signal $Y(k) = -Z(2k-1)$ is

$$Y(k) = 2e^{-j2\pi(3(2k-1)-15)/32} = 2e^{-j2\pi(3k-9)/16}.$$

It is periodic with period $N/Q = 16$. Group delay of this signal is $n_0 = 3$ with period 16. Therefore within $n = 0, 1, \dots, 31$ group delays $n_0 = 3$ and $n_0 + 16 = 19$ of $Y(k)$ correspond to the missing sample positions. The signal must have the form $X(k_{0m}) \in \{2e^{-j2\pi(3k-9)/16} \mid k = 0, 1, \dots, \frac{N}{Q} - 1\}$, with $k = 3$ corresponding to $k_{0m} = 2k - 1 = 5$ producing $X(5) = 2$.

V. CONCLUSION

Reconstruction of a sparse signal, using recently proposed gradient-based method, is done by considering missing samples as variables. Theorems for the uniqueness of the solution obtained by varying missing samples, in the case of an arbitrary and already reconstructed signal, are stated and proved.

The calculation complexity of the proposed theorems is low. The theory is illustrated on numerical and statistical examples.

Proposed approach can be extended to other signal transforms including nonredundant basis (dictionaries). One of the possible redundant basis closely related to the presented DFT is short time Fourier transform with overlapping windows [18] [19].

REFERENCES

- [1] I. Djurović, L. Stanković, J. F. Böhme, "Robust L-estimation based forms of signal transforms and time-frequency representations", *IEEE Trans. on SP*, vol. 51, no. 7, pp. 1753-1761, 2003.
- [2] L. Stanković, I. Orović, S. Stanković, and M. Amin, "Robust Time Frequency Analysis based on the L-estimation and Compressive Sensing," *IEEE Signal Processing Letters*, vol. 20, no. 5, pp. 499-502, 2013.
- [3] D. L. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289-1306, 2006.
- [4] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Information Theory*, vol. 52, no. 2, pp. 489-509, 2006.
- [5] E. J. Candès, "The restricted isometry property and its implications for compressed sensing", *Comptes Rendus Mathématique*, vol. 346, no. 9, pp. 589-592, 2008.
- [6] J. Ma, X. Yuan, and L. Ping, "Turbo Compressed Sensing with Partial DFT Sensing Matrix", *IEEE Signal Processing Letters*, vol. 22, no. 2, pp. 158-161, 2015.
- [7] A. Tulino, G. Caire, S. Verdu, and S. Shamai, "Support recovery with sparsely sampled free random matrices," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4243-4271, Jul. 2013.
- [8] S. Foucart, H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Springer Birkhauser, New York, 2013
- [9] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 20, no. 3, pp. 301-321, 2009.
- [10] 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.
- [11] G. Davis, S. Mallat and M. Avellaneda, "Greedy adaptive approximation," *Journal of Constructive Approximation*, vol. 12, pp. 57-98, 1997.
- [12] P. Flandrin and P. Borgnat, "Time-Frequency Energy Distributions Meet Compressed Sensing," *IEEE Trans. on Signal Processing*, vol. 58, no. 6, 2010, pp. 2974-2982.
- [13] J. A. Tropp, J.N. Laska, M.F. Duarte, J.K. Romberg, R.G. Baraniuk, "Beyond Nyquist: Efficient Sampling of Sparse Bandlimited Signals," *Information Theory, IEEE Transactions on*, vol.56, no.1, pp.520,544, Jan. 2010.
- [14] H. Tsutsu, Y. Morikawa, "An lp Norm Minimization Using Auxiliary Function for Compressed Sensing", in *Proc. IMECS*, March 2012, Hong Kong.
- [15] S. Aviyente, "Compressed Sensing Framework for EEG Compression," in *Proc. Stat. Sig. Processing*, 2007, Aug. 2007.
- [16] 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.
- [17] L. Stanković, M. Daković and S. Vujović, "Adaptive Variable Step Algorithm for Missing Samples Recovery in Sparse Signals," *IET Signal Processing*, vol. 8, no. 3, pp. 246 -256, 2014, (arXiv:1309.5749v1).
- [18] S. G. Mallat, Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397-3415, 1993
- [19] LJ. Stanković, "On the STFT Inversion Redundancy," *IEEE Transactions on Circuits and Systems II*, DOI: 10.1109/TCSII.2015.2482438, in print, 2015.