Fast communication

# Missing samples analysis in signals for applications to L-estimation and compressive sensing

Ljubisa Stankovic [a],[*], Srdjan Stankovic [a], Moeness Amin [b]

[a] University of Montenegro, Podgorica, Montenegro
[b] Villanova University, Villanova, USA

## ARTICLE INFO

## ABSTRACT

This paper provides statistical analysis for efficient detection of signal components when missing data samples are present. This analysis is important for both the areas of L-statistics and compressive sensing. In both cases, few samples are available due to either noisy sample elimination or random undersampling signal strategies. The analysis enables the determination of the sufficient number of observation and as such the minimum number of missing samples which still allow proper signal detection. Both single component and multicomponent signals are considered. The results are verified by computer simulations using different component frequencies and under various missing-available samples scenarios.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

Robust transforms have been introduced to deal with signals affected by impulsive noise [1–18]. These transforms, which use robust statistics and Huber's estimation theory, were originally applied to the Fourier transform and then extended to the time–frequency analysis [1,2]. One of the most effective robust methods is the L-estimation form, obtained by using the L-statistics. The main idea is to remove samples corrupted by impulsive noise (by using trimming procedure) and to compute the transform using the remaining incomplete set of samples. Thus, the L-estimation can lead to significantly reduced number of signal samples or observations. Having in mind random positions of these observations, the problem can be cast as compressive sensing (CS) if the underlying signal is sparse in a particular domain. In this

respect the task becomes sparse signal reconstruction with a large number of missing samples [7–17]. Although there are numerous papers on the L-estimation, the appropriate number of samples that should be omitted remained unclear [2]. As a consequence, we might end up throwing much more samples than needed to ensure noise elimination. The missing samples effect as a source of generation of undesirable sidelobes in the spectral analysis of non-uniformly sampled data sequences has been considered in [19]. In this paper, analysis of the effects of missing samples on transform values is performed from a noise perspective. It is shown that the missing samples, when using the discrete Fourier transform (DFT), could be described by a new type of noise which deteriorates signal representation [20]. Increasing the number of missing samples increases the corresponding noise level, rendering signal component detection more difficult. The expression which relates the number of missing samples to the statistics of this noise is derived. This relation is also crucial for the analysis of the initial steps (transforms) in the sparse signal reconstruction algorithms [7–13].

* Corresponding author. Tel.: +38 267234430; fax: +38 220244921.
E-mail addresses: l.stankovic@ieee.org, ljubisa@ac.me (L. Stankovic),
srdjan@ac.me (S. Stankovic), moeness.amin@villanova.edu (M. Amin).

The paper is structured as follows. After Introduction, the variance of induced noise, generated by missing samples, is derived. This analysis is further extended to general sparse signals. Theoretical results are illustrated by numerical examples.

## 2. Theory

Both the L-estimation based robust signal analysis and the CS deal with an incomplete set of randomly selected samples. Unlike in the L-estimation where the reduced observations are the consequence of omitting corrupted noisy samples, in the CS the missing observations are the result of the sampling strategy. In both cases, the signal is assumed sparse in the transformation domain.

Consider a set of $M$ signal values:

$$\Theta = \{s(1), s(2), \ldots, s(M)\}.$$

Without loss of generality, assume that the signal samples are zero-mean. Furthermore, consider the DFT domain, as an example of a sparse signal transform and a study case in this paper. A signal which is sparse in the DFT domain can be written as

$$s(n) = \sum_{i=1}^{K} A_i e^{j2\pi k_{0i} n/M}, \tag{1}$$

where the level of sparsity is $K \ll M$, while $A_i$ and $k_{0i}$ denote amplitudes and frequencies of the signal components, respectively. The DFT of this signal is

$$S(k) = \sum_{n=1}^{M} \sum_{i=1}^{K} A_i e^{-j2\pi(k-k_{0i})n/M} = \sum_{n=1}^{M} \sum_{i=1}^{K} A_i x(n). \tag{2}$$

Normalized signal components $\exp(j2\pi n k_{0i}/M)$ are multiplied by the basis function $\exp(-j2\pi n k/M)$ to produce $x(n) = \exp(-j2\pi n l_i/M)$, where $l_i = k - k_{0i}$ is assumed to be an integer. Values of $x(n)$ are from the set

$$\Phi = \{e^{-j2\pi n l_i/M}, \ n = 1, 2, \ldots, M\}. \tag{3}$$

Note that $k \neq k_{0i}$ implies $l_i \neq 0$.

Relation among the members of set $\Phi$ is

$$x(1) + x(2) + \cdots + x(M) = 0. \tag{4}$$

This is evident for any $l_i \neq 0$. Cases when $l_i = 0$ will be analyzed separately.

Consider a subset of $M_A \leq M$ randomly positioned available samples from $\Phi$:

$$\Psi = \{y(1), y(2), \ldots, y(M_A)\} \subset \Phi. $$

This set corresponds to a compressed signal. The same holds after applying the L-statistics to $x(n)$ to remove the corrupted samples. In both cases, $M_Q = M - M_A$ signal samples are unavailable though for different reasons.

In linear discrete signal transforms, where the inner products are performed between the signal values and the basis functions, omitting some of the signal samples produces the same result as if these samples assume zero values. Thus, the basic idea is to model missing samples with zero values, since any sum over all values in the subset $\Psi = \{y(1), y(2), \ldots, y(M_A)\}$ will be the same as the sum over all values within the complete set $\Phi = \{x(1), x(2), \ldots, x(M)\}$ with removed/unavailable samples, belonging to $\Upsilon = \Phi - \Psi$, being set to zero.

A transform with a reduced number of signal samples can be considered as transform of complete set of samples, affected by disturbance (noise). For $K = 1$ and $A_1 = 1$, $k_{01} = k_0$, $l = k - k_0$,

$$\varepsilon(n) = \begin{cases} 0 & \text{for remaining signal samples} \\ -x(n) = -e^{j2\pi n l/N} & \text{for removed (unavailable) signal samples}. \end{cases}$$

The DFT over the available set of samples from $\Psi$ will be

$$X(k) = X_l = \sum_{n=1}^{M_A} y(n) = \sum_{n=1}^{M} [x(n) + \varepsilon(n)].$$

It is a random variable formed as a sum of $M_A$ randomly positioned samples

$$y(n) \in \Psi \subset \Phi = \{x(1), x(2), \ldots, x(M)\}.$$

In our example, this random variable corresponds to the DFT of a CS signal or to the L-estimation based DFT calculation. Samples $x(n)$ depend on $k$. The same holds for $y(n)$. Note that due to (4), the statistics is the same for the sum of complementary samples (disturbance $\varepsilon(n)$) from the set $\Upsilon = \Phi - \Psi$ containing $M_Q = M - M_A$ missing values.

Obviously, for $l \neq 0$,

$$E\{X_{l \neq 0}\} = 0$$

since $E\{y(n)\} = 0$ and $E\{x(m)\} = 0$. For $l = 0$ ($k = k_0$)

$$E\{X_{l=0}\} = M_A$$

holds, since for this value of $l$ all $x(n) = 1$. Thus, $X$ is a zero mean random variable for $l \neq 0$ and deterministic for $l = 0$. Therefore, removing/omitting samples in the L-estimation or in the initial transform of the CS algorithm corresponds to a new additive noise. The resulting noise depends on the signal, the total number of samples, and the number of missing samples.

Let us now calculate the variance of this noise in $X_l$ for $l \neq 0$. It is defined by

$$\begin{aligned} var\{X\} &= E\{[y(1) + y(2) + \cdots + y(M_A)][y(1) + y(2) + \cdots + y(M_A)]^*\} \\ &= E\{|y(1)|^2\} + E\{|y(2)|^2\} + \cdots + E\{|y(M_A)|^2\} \\ &\quad + E\{y(1)y^*(2)\} + E\{y(1)y^*(3)\} + \cdots + E\{y(1)y^*(M_A)\} \\ &\quad + E\{y(2)y^*(1)\} + E\{y(2)y^*(3)\} + \cdots + E\{y(2)y^*(M_A)\} \\ &\quad \vdots \\ &\quad + E\{y(M_A)y^*(1)\} + E\{y(M_A)y^*(2)\} \\ &\quad + \cdots + E\{y(M_A)y^*(M_A - 1)\} \end{aligned} \tag{5}$$

Obviously,

$$E\{|y(1)|^2\} + E\{|y(2)|^2\} + \cdots + E\{|y(M_A)|^2\} = M_A.$$

According to the assumption, all signal samples are not statistically independent for $l \neq 0$. They satisfy (4). By multiplying (4) with $x(i)$ we obtain

$$x(i)[x(1) + x(2) + \cdots + x(M)]^* = 0$$

with expectation

$$\sum_{j=1}^{M} E\{x(i)x^*(j)\} = 0. \tag{6}$$

Since all variables $x(n)$ are equally distributed we may write

$E\{x(i)x^*(j)\} = B$   for $i \neq j$,
$E\{x(i)x^*(i)\} = 1$.

From (6) we obtain

$$B = E\{x(i)x^*(j)\} = E\{y(i)y^*(j)\} = -\frac{1}{M-1} \quad \text{for } i \neq j.$$

Now we can easily calculate terms in (5), for example, the first line (with mixed terms) produces

$E\{y(1)y^*(2)\} + E\{y(1)y^*(3)\} + \cdots$
$\quad + E\{y(1)y^*(M_A)\} = (M_A - 1)\left(-\frac{1}{M-1}\right).$

The same result holds for all other lines with mixed terms in (5). Finally, the variance of $X_l$, for $l \neq 0$, is

$$var\{X_{l \neq 0}\} = M_A + (M_A - 1)\left(-\frac{1}{M-1}\right)M_A = M_A \frac{M - M_A}{M-1}. \quad (7)$$

For $M_A \ll M$ a rough approximation $var\{X_{l \neq 0}\} \approx M_A$ follows. It corresponds to the assumption of statistically independent values $x(n)$.

## 3. Amplitude analysis of the signal and CS noise in the DFT

It is easy to find the ratio of the DFT value at the signal position $l = 0$ and the DFT value at any other position $l \neq 0$. This ratio may be used as a measure of the wrong signal detection event. The DFT value at the signal position is $X_{l=0} = M_A$ since all terms are added in phase for the signal component. Variable $X_{l \neq 0}$ is complex-valued, zero-mean, with variance (7). If the number of missing or available samples is not too small, then according to the central limit theorem the real and imaginary parts of variable $X_{l \neq 0}$ behave as Gaussian variables with variance

$$\frac{\sigma^2}{2} = \frac{var\{X_{l \neq 0}\}}{2} = M_A \frac{M - M_A}{M-1}/2.$$

Absolute value of $X_{l \neq 0}$ is therefore Rayleigh distributed. In the Rayleigh distribution, a random variable $|X_{l \neq 0}|$ is below $T_{95} = \sqrt{3}\sigma$ with probability 0.9508. Thus, the ratio

$$\left|\frac{X_{l \neq 0}}{X_{l=0}}\right| < \frac{\sqrt{3}\sigma}{M_A} = \sqrt{\frac{3(M - M_A)}{M_A(M-1)}} = R_{95} \quad (8)$$

holds with probability 0.95. For example, for $M = 256$ and $M_A = 32$, the value of $|X_{l \neq 0}|$ is greater than $0.27|X_{l=0}|$ with probability 5%. Threshold for any other probability can be easily calculated.

This ratio can be used to find the range of $M_A$ that the initial CS and the L-estimate algorithm will be able to detect signal component. It is especially important in multicomponent signal cases, whose detailed analysis follows. Then, all components, for a given $M_Q = M - M_A$, below the noise level will be lost in the DFT. The ratio $R_{95}$ as a function of $M_Q$ is presented in Fig. 1. For small number of missing samples $M_Q$ (large number of available samples $M_A$), this ratio is very small. It means that the values $X_{l \neq 0}$, caused by missing samples, are very small compared to the signal value $X_{l=0}$. On the other side, just a few available samples $M_A = M - M_Q$ are enough to make this ratio lower than 1 and as such enable detection of a sinusoid in one-component signal. This is in accordance with the fact that a real sinusoidal signal is determined by only three independent signal samples [21].

Before we proceed with multicomponent signals, the theoretical results from this section are statistically verified. Two quantities are calculated and checked. The variance is calculated according to relation (7) for various $M$ and $M_A$. Then, the statistical results are obtained for the same variance and various $M$ and $M_A$ based on 100 000 realizations with randomly positioned samples. The results for variance produced by the theory are given in Table 1, in the column denoted by Var(T). The statistical results for the same variance are presented as Var(E) in the same table. It is obvious that the results matching is very high.

The level $R_{95}$ is also calculated using (8) for various $M$ and $M_A$. Then the DFT with removed/unavailable samples is calculated. The DFT values at $l \neq 0$ above $R_{95}|X_{l=0}|$ are detected and their percent (with respect to the total number of DFT value) is calculated. It should be just below 5%. The statistically obtained percent for various $M$ and $M_A$ is given in the column denoted by $P_{5\%}$ of Table 1. Matching with the true value (4.98%) is also high.

*Case*1: For $M = 128$ samples and various numbers of randomly positioned missing samples, the variance relation (7) and the level of 0.95 are calculated. The results are given in Table 1 (left). Note that the variance values (Var) for $M_Q > M/2$ are the same as for $M - M_Q$. This is obvious
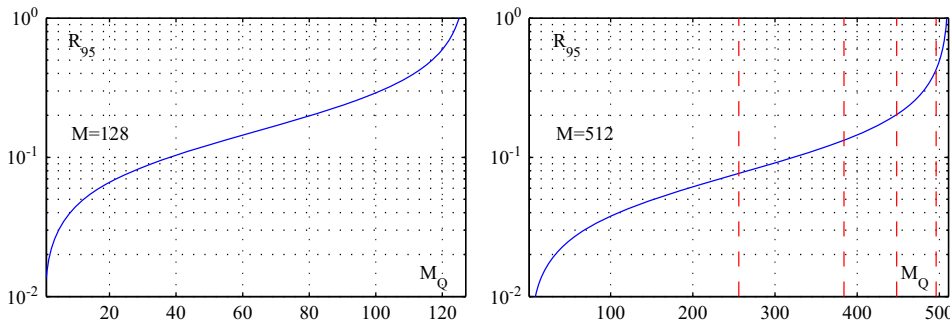


**Fig. 1.** The DFT amplitude ratio at the signal position and outside the signal position for $M = 128$ (left) and $M = 512$ (right). The values of $R_{95}$ for $M = 512$ and $M_Q = 256, 384, 448$, and 496 are indicated by dashed red lines. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

from (7). It is statistically verified and presented in Table 1 (for $M_Q = M-32 = 96$ and $M_Q = M-16 = 112$).

*Case*2: The same results are presented for $M=512$ and $M_Q$ up to $M/2$. For $M_Q > M/2$ the results are the same as for $M-M_Q$, Table 1 (right).

The illustration of the DFT for a single signal realization, with various number of available samples, is presented in Fig. 2. The reference level of $R_{95}|X_{l=0}|$, calculated according to (8), is presented by a horizontal dashed line. In multicomponent signals, a similar reference level will play a role of a detection threshold. Note also that if the noise-alone values of $X(k)$ satisfy the condition that they are approximately zero-mean Gaussian, then it can be checked that they are also uncorrelated and approximately jointly Gaussian for various values of $k$.

If a window is used in the DFT calculation, then for some common window forms the windowed DFT values can be easily related to the ones calculated without a window. For example, for the Hann(ing) window over the whole data set the windowed DFT is $X_H(k) = 0.25X(k-1) + 0.5X(k) + 0.25X(k + 1)$. The variance of $X_H(k)$ with the Hann(ing) window is $\sigma_H^2 = 3\sigma^2/8$. Similar relations can be written for the Hamming or the Blackman window.

## 4. Multicomponent signal analysis

The analysis will be extended to a variable $x(n) \in \Phi$ of the form

$$x(n) = A_1 \exp(-j2\pi nl_1/M) + A_2 \exp(-j2\pi nl_2/M)$$
$$+ \cdots + A_K \exp(-j2\pi nl_K/M),$$

corresponding to a $K$-component signal, defined by (1) and (2), where $l_1 = k-k_{01}$, $l_2 = k-k_{02}$, ..., $l_K = k-k_{0K}$ in the study case of DFT analysis. Values of $x(n)$ depend on $k$ as well. The mean value of a subset $\Psi$ of $M_A \leq M$ randomly positioned values $y(n) \in \Psi \subset \Phi = \{x(1), x(2), ..., x(M)\}$ is

$$E\{X_l\} = A_1 M_A \delta(l_1) + A_2 M_A \delta(l_2) + \cdots + A_K M_A \delta(l_K).$$

The DFT variance at the points where there are no signal components (i.e., $l_1 \neq 0$, $l_2 \neq 0$, ... $l_K \neq 0$) is

$$\sigma_N^2 = var\{X_{l_i \neq 0}\} = \sum_{j=1}^K A_j^2 M_A \frac{M-M_A}{M-1}, \qquad (9)$$

since the signal components are uncorrelated zero-mean. At a frequency point of a signal component, for $l_i = 0$, we get

$$\sigma_{S_i}^2 = var\{X_{l_i = 0}\} = \sum_{\substack{j=1 \\ j \neq i}}^K A_j^2 M_A \frac{M-M_A}{M-1} \qquad (10)$$

for the DFT variance. It means that a DFT of the signal sample $X_{l_i = 0}$ will be a random variable disturbed by an additive noise so that the resulting variance is $\sigma_{S_i}^2$, while

**Table 1**
Statistical variance (denoted by (E)) with theory obtained one (denoted by (T)) for $M=128$ (left) and $M=512$ (right) and various number of missing samples. The probability level of $4.98\% \approx 5\%$ is statistically checked.

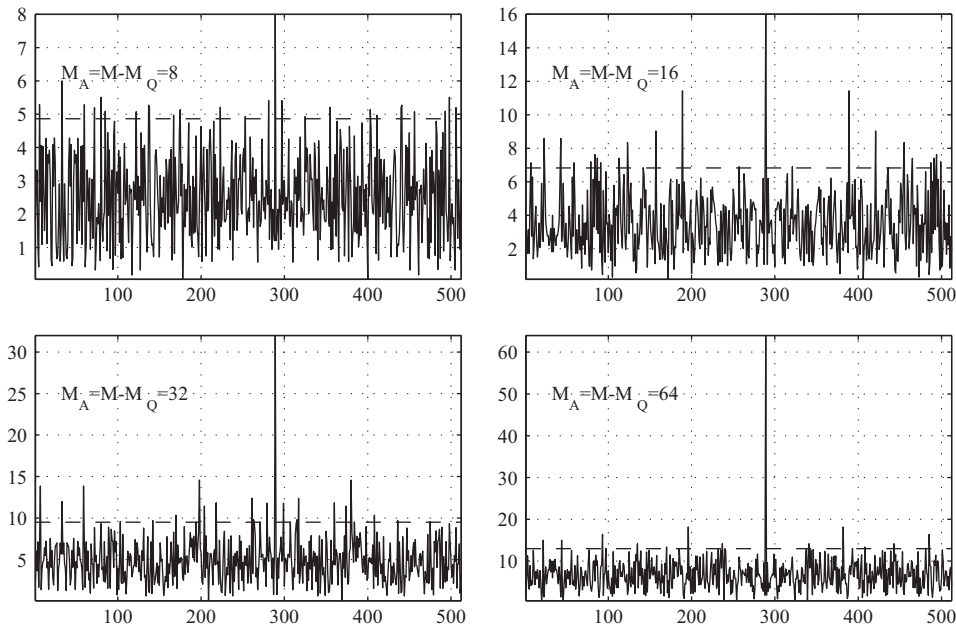| $M_Q$ | $P_{5\%}$ | Var(E) | Var(T) | $M_Q$ | $P_{5\%}$ | Var(E) | Var(T) |
|---|---|---|---|---|---|---|---|
| 8 | 4.64 | 7.60 | 7.56 | 16 | 4.87 | 15.55 | 15.53 |
| 16 | 4.81 | 14.16 | 14.11 | 32 | 4.80 | 29.85 | 30.06 |
| 24 | 4.76 | 19.64 | 19.65 | 48 | 4.90 | 43.71 | 43.59 |
| 32 | 4.79 | 24.15 | 24.19 | 64 | 4.95 | 56.04 | 56.11 |
| 40 | 4.89 | 27.65 | 27.72 | 96 | 5.01 | 78.29 | 78.15 |
| 48 | 4.97 | 30.26 | 30.24 | 128 | 4.89 | 95.96 | 96.19 |
| 56 | 4.91 | 31.71 | 31.74 | 160 | 5.01 | 110.14 | 110.22 |
| 64 | 4.93 | 32.15 | 32.25 | 192 | 4.95 | 120.17 | 120.23 |
| 96(32) | 4.77 | 24.13 | 24.19 | 224 | 4.91 | 126.12 | 126.25 |
| 112(16) | 4.80 | 14.19 | 14.11 | 256 | 4.93 | 127.69 | 128.25 |



**Fig. 2.** Illustration of the DFT of one signal component for different number of available samples along with the calculated threshold level for 95% of samples obtained by theory. Threshold level $R_{95}M_A$ is plotted by a horizontal dashed line.

the noise only DFT values will be random variables with variance $\sigma_N^2$.

Next, we will find a probability that a DFT value of noise at any $l_1 \neq 0$, $l_2 \neq 0$, … $l_K \neq 0$ is higher than a signal DFT value, for example, at $l_1 \neq 0$, …, $l_{i-1} \neq 0$, $l_i = 0$, $l_{i+1} \neq 0$,…, $l_K \neq 0$. This case corresponds to a false signal detection of the $i$th component.

Real and imaginary parts of noise-alone DFT value can be described by Gaussian distribution, according to the central limit theorem:

$$\mathcal{N}(0, \sigma_N^2/2)$$

with zero mean and variance $\sigma_N^2$ defined by (9). Real and imaginary parts of the signal DFT value can be described by Gaussian distribution

$$\mathcal{N}(M_A A_i, \sigma_{S_i}^2/2), \quad \mathcal{N}(0, \sigma_{S_i}^2/2), \tag{11}$$

respectively, where $M_A A_i$ is the mean and variance $\sigma_{S_i}^2$ is defined by (10). A real-valued $A_i$ is assumed without any loss of generality.

The probability density function (pdf) for the absolute DFT values at the position of the $i$th signal component, $l_i = 0$ (whose real and imaginary parts are described by (11)) is Rice-distributed:

$$p(\xi) = \frac{2\xi}{\sigma_{S_i}^2} e^{-(\xi^2 + A_i^2 M_A^2)/\sigma_{S_i}^2} I_0\left(A_i M_A 2\xi/\sigma_{S_i}^2\right), \quad \xi \geq 0 \tag{12}$$

where $I_0$ is the zero-order modified Bessel function.

The probability density function for the absolute DFT values outside the signal components is Rayleigh-distributed (Rice-distribution with $A_i M_A = 0$ and $I_0(0) = 1$)

$$q(\xi) = \frac{2\xi}{\sigma_N^2} e^{-\xi^2/\sigma_N^2}, \quad \xi \geq 0. $$

The DFT at a noise-alone position takes a value greater than $\varXi$, with probability

$$Q(\varXi) = \int_{\varXi}^{\infty} \frac{2\xi}{\sigma_N^2} e^{-\xi^2/\sigma_N^2} \, d\xi = \exp\left(-\frac{\varXi^2}{\sigma_N^2}\right). \tag{13}$$

The probability that a DFT of noise-alone is lower than $\varXi$ is $[1 - Q(\varXi)]$. The total number of noise-alone points is $M_K = M - K$, where $K$ is the number of signal points. The probability that $M_K$ independent DFT noise-alone values are lower than $\varXi$ is $[1 - Q(\varXi)]^{M_K}$. Probability that at least one of $M_K$ DFT noise-alone values is greater than $\varXi$ is

$$G(\varXi) = 1 - [1 - Q(\varXi)]^{M_K}. \tag{14}$$

When a noise-alone DFT value surpasses the DFT signal value, then an error in the considered methods occurs. To calculate this probability, consider the absolute DFT value of a signal component at and around $\xi$. The DFT value at the signal position is within $\xi$ and $\xi + d\xi$ with the probability $p(\xi)d\xi$, where $p(\xi)$ is defined by (12). The probability that at least one of $M_K$ DFT noise-alone values is above $\xi$ in amplitude is $G(\xi) = 1 - [1 - Q(\xi)]^{M_K}$. Thus, the probability that the absolute DFT signal component value is within $\xi$ and $\xi + d\xi$ and that at least one of the absolute DFT noise-alone values outside the DFT signal value exceeds the DFT signal value is $G(\xi)p(\xi)d\xi$. Considering all possible values of $\xi$, from (13) and (14), it follows that the probability of the wrong detection of the $i$th signal

component is

$$P_E = \int_0^{\infty} G(\xi)p(\xi) \, d\xi$$

$$= \int_0^{\infty} \left(1 - \left[1 - \exp\left(-\frac{\xi^2}{\sigma_N^2}\right)\right]^{M_K}\right) \frac{2\xi}{\sigma_{S_i}^2} e^{-(\xi^2 + A_i^2 M_A^2)/\sigma_{S_i}^2} I_0$$

$$\times (A_i M_A 2\xi/\sigma_{S_i}^2) \, d\xi. \tag{15}$$

### 4.1. Approximative error expression

Approximation of this expression can be calculated by assuming that the DFT of the $i$th signal component is not random and that it is equal to $M_A A_i$ (positioned at the mean value of the signals DFT). This approximation assumes that the influence of noise to amplitude is symmetric and equally increases and decreases the DFT signal value. The form of error probability is then very simple

$$P_E \cong 1 - \left[1 - \exp\left(-\frac{M_A^2 A_i^2}{\sigma_N^2}\right)\right]^{M_K}. \tag{16}$$

This expression can be used for simple rough approximative analysis.

The mean of a Rice-distributed variable described by (12) is

$$M_{Rice} = \sigma_{S_i} \sqrt{\pi} L_{1/2}(-M_A^2 A_i^2/\sigma_{S_i}^2)/2 \cong M_A A_i,$$

where $L_{1/2}(x) = e^{x/2}[(1-x)I_0(-x/2) - xI_1(-x/2)]$, and $I_1(x)$ is the modified first-order Bessel function. The variance of the Rice-distributed variable is

$$var_{Rice} = \sigma_{S_i}^2 + M_A^2 A_i^2 - M_{Rice}^2 \cong \sigma_{S_i}^2.$$

From numerical analysis, we concluded that a closer approximation than (16) is achieved if the Rice distribution mean is slightly corrected, for one standard deviation, as $M_{Rice} - \sqrt{var_{Rice}}$. The explanation lies in the fact that the DFT amplitudes lower than the mean value contribute more to the error than those above the mean value. Then,

$$P_E \cong 1 - \left[1 - \exp\left(-\frac{(M_A A_i - \sigma_{S_i})^2}{\sigma_N^2}\right)\right]^{M_K}. \tag{17}$$

It will be shown that this simple expression is a good approximation. It predicts almost exactly 1% error level as the very complex integral form in (15). The variances $\sigma_N^2$ and $\sigma_{S_i}^2$ are defined by (9) and (10). Note that as far as $M$ is large and the signal is sparse ($K \ll M$) the exponent in (16) and (17) is large. Then, the exponent variations from $M - K$ to $M$ will have small influence to the expression value. Thus, for $M$ and $K$ satisfying the above conditions we may use $M$ instead of $M_K$ in the approximative error expressions.

The approximations will be checked on an example with a sparse signal. Probabilities of detection error for a four component signal of amplitudes $A_1 = 1$, $A_2 = 0.5$, $A_3 = 0.25$, and $A_4 = 0.1$ are shown in Fig. 3. Probabilities are calculated for each signal component, according to (15) and (17). Results obtained by the exact integral (15) are
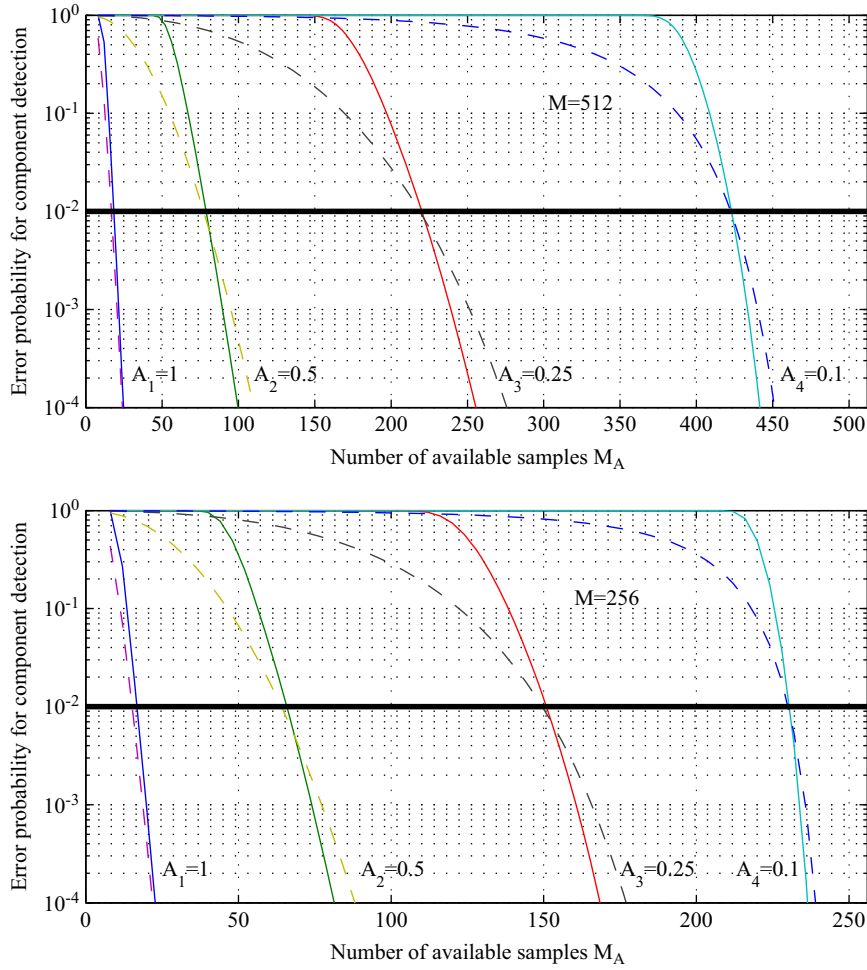
**Fig. 3.** Probability that a component will not be detected in a four components signal as a function of the number of available samples $M_A$ (full lines approximative expression and dashed line exact integral expression for the error). Thick line is the 1% error probability line.

shown by a dashed line, while the ones calculated using the simple approximation (17) are given by a solid line. We can see that the strongest component is detectable with a very small number of available samples, as expected. The number of required available samples increases as the amplitude of a component decreases. A level of error probability of 1% is presented by a thick horizontal line in this figure. Both results are highly consistent at this important level, which defines the number of samples required to detect a component with 99% probability.

Now we will discuss the results on a single realization of the considered sparse signal, for $M = 512$ and various number of available samples, Fig. 4.

Based on the theoretically calculated level, it is easy to conclude that: (1) The strongest signal component with $A_1 = 1$ will be detectable with a quite small number of random samples. From Fig. 3 ($M = 512$, upper subplot) we see that $M_A > 15$ random samples are sufficient. (2) From the same figure we see that the next component with $A_2 = 0.5$ will be detected with $M_A > 75$. (3) The component

with $A_3 = 0.25$, in this case, requires $M_A > 220$ random samples, while (4) for the weakest component with $A_4 = 0.1$ we need $M_A > 420$ samples in order to detect it. This is in accordance with the single DFT realization shown in Fig. 4 for different number of the available samples $M_A$. For the first two subplots ($M_A = 16$ and $M_A = 32$) only the strongest component is detectable in the L-estimation or in the initial transform in the CS. For $M_A = 128$ two components are detected in this stage, while $M_A = 384$ clearly indicates existence of three components in the analyzed signal with missing $M_Q = M - M_A = 128$ samples. The 5% level is defined by the threshold $T_{95} = \sqrt{3}\sigma_N$ with $\sigma_N^2 = A_e^2 M_A(M - M_A)/(M-1)$. The effective signal value $A_e$ can be estimated using $M_A$ available signal samples $s(n)$ as $A_e^2 = \sum_{j=1}^K A_j^2 = \sum_n |s(n)|^2/M_A$. In addition to the 5% level here we have also drawn the 0.1% level, calculated using $T_{99.9} = 1.5T_{95}$.

The results presented in this paper can be used to improve performance and to reduce calculation complexity of the standard CS algorithms. Research in this direction is under the way. Recently, significant efforts have
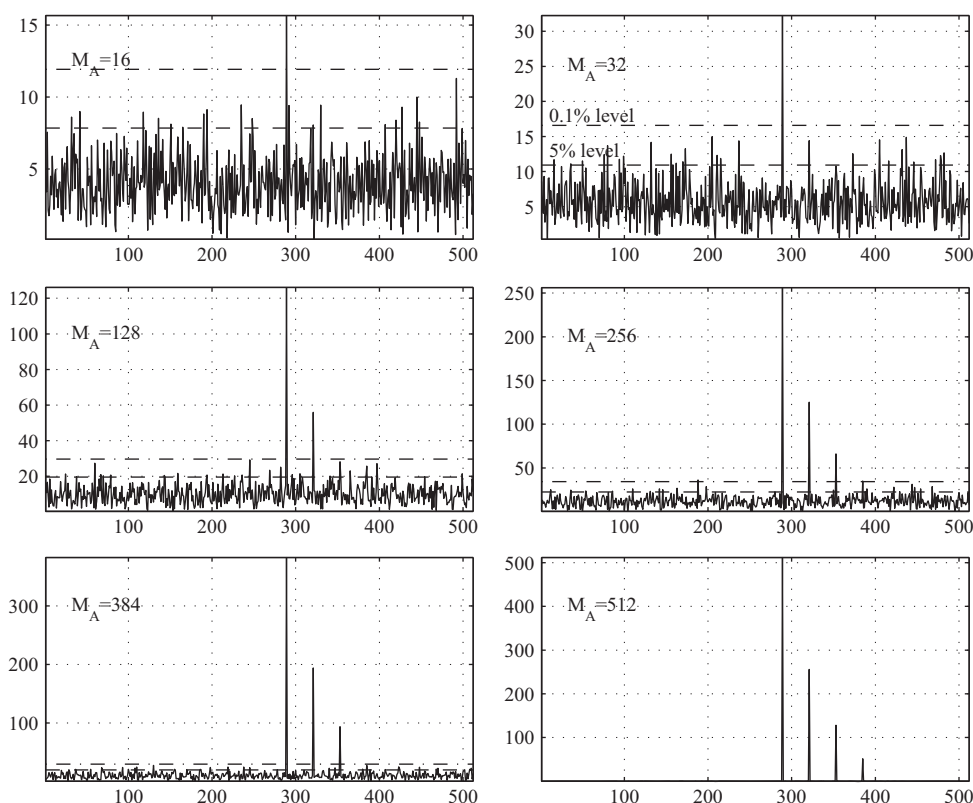
**Fig. 4.** Illustration of four component signal for different number of available samples, along with the threshold calculated by theory. Detection threshold $T_{95}$ for 5% level is presented by a dash line, while the threshold $T_{99}$ for 0.1% level is given by a dash-dot line.

been done to extend the notion of sparsity and to apply the CS algorithms to the signals that are not sparse in any common discrete transformation domain. Such an example is the spectral analysis of close sinusoids with frequencies outside the frequency grid. They are not sparse in the DFT, but could be processed using a combination of the spectral domain interpolation (to improve sparsity) and the high-resolution methods, like the MUSIC with the CS algorithms [22]. Extension of the presented theory beyond the standard sparse signals framework to the results that could be obtained by combinations of several methods and algorithms is a significant challenge for future research.

## 5. Conclusion

We provided analysis for the determination of the number of missing samples allowed for accurate detection of sinusoidal signals. This analysis proves useful when applying L-statistics and when using sparse signal reconstruction techniques with Fourier basis. The latter is clearly the case in orthogonal matching pursuit. The analytical expressions derived are validated by computer simulations for a single and multiple components and under different numbers of missing and available samples. The proposed approach can be generalized for all existing L-estimate based signal transforms and L-estimation based time–frequency distributions.

## References

[1] V. Katkovnik, Robust M-periodogram, IEEE Transactions on Signal Processing 46 (November (11)) (1998) 3104–3109.

[2] I. Djurović, L. Stanković, J.F. Böhme, Robust L-estimation based forms of signal transforms and time–frequency representations, IEEE Transactions on Signal Processing 51 (7) (2003) 1753–1761.

[3] R.E. Carrillo, K.E. Barner, 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 4 (2) (2010) 392–408.

[4] B. Barkat, L. Stankovic, Analysis of polynomial FM signals corrupted by heavy-tailed noise, Signal Processing 84 (2004).

[5] A.A. Roenko, V.V. Lukin, I. Djurovic, Two approaches to adaptation of sample myriad to characteristics of SalphaS distribution data, Signal Processing 90 (7) (2010) 2113–2123.

[6] J.L. Paredes, G.R. Arce, Compressive sensing signal reconstruction by weighted median regression estimates, IEEE Transactions on Signal Processing 59 (6) (2011) 2260–2585.

[7] D. Donoho, Compressed sensing, IEEE Transactions on Information Theory 52 (4) (2006) 1289–1306.

[8] T. Strohmer, Measure what should be measured: progress and challenges in compressive sensing, IEEE Signal Processing Letters 19 (December (12)) (2012) 887–893.

[9] E. Sejdic, A. Can, L.F. Chaparro, C.M. Steele, T. Chau, Compressive sampling of swallowing accelerometry signals using time–frequency dictionaries based on modulated discrete prolate spheroidal sequences, EURASIP Journal on Advances in Signal Processing, 2012 101 (2012).

[10] J.H.G. Ender, On compressive sensing applied to radar, Signal Processing 90 (May (5)) (2010) 1402–1414.

[11] S. Kunis, H. Rauhut, Random sampling of sparse trigonometric polynomials, II. Orthogonal matching pursuit versus basis pursuit, Journal Foundations of Computational Mathematics Archive 8 (November (6)) (2008) 737–763.

[12] H. Rauhut, Stability results for random sampling of sparse trigonometric polynomials, IEEE Transactions on Information Theory 54 (December (12)) (2008) 5661–5670.

[13] S. Aviyente, Compressed sensing framework for EEG compression, in: Proceedings of Statistical Signal Processing, 2007, August 2007.

[14] K.O. Egiazarian, A. Foi, V. Katkovnik, Compressed sensing image reconstruction via recursive spatially adaptive filtering, in: Proceedings of ICIP, 2007.

[15] X. Zeng, L. Yang, Mixed impulse and Gaussian noise removal using detail-preserving regularization, Optical Engineering 49 (9) (2010).

[16] P. Flandrin, P. Borgnat, Time–frequency energy distributions meet compressed sensing, IEEE Transactions on Signal Processing 58 (6) (2010) 2974–2982.

[17] J. Meng, J. Ahmadi-Shokouh, H. Li, Z. Han, S. Noghanian, E. Hossain, Sampling rate reduction for 60 GHz UWB communication using compressive sensing, in: Asilomar Conference on SSSC, Monterey, CA, 2009.

[18] Y. Wang, P. Stoica, J. Li, T.L. Marzetta, Nonparametric spectral analysis with missing data via the EM algorithm, Digital Signal Processing March (2005).

[19] P. Stoica, J. Li, H. He, Spectral analysis of nonuniformly sampled data: a new approach versus the periodogram, IEEE Transactions on Signal Processing 57 (March) (2009) 843–858.

[20] L. Stanković, S. Stanković, I. Orović, M. Amin, Robust time–frequency analysis based on the L-estimation and compressive sensing, IEEE Signal Processing Letters 20 (5) (2013) 499–502.

[21] L. Stanković, M. Daković, T. Thayaparan, Time–Frequency Signal Analysis with Applications, Artech House, 2013.

[22] M.F. Duarte, R.G. Baraniuk, Spectral compressive sensing, Applied and Computational Harmonic Analysis 35 (1) (2013) 111–129.