Algorithms of signal parameter estimation based on the theory of Compressive Sensing and their computer simulation

V I Parfenov¹, D Y Golovanov² and N A Kunaeva³

¹Department of Physics, Voronezh State University, 1 Universitetskaya pl., Voronezh, 394018, Russian Federation
²Department of Applied Mathematics, Computational Science and Mechanics, Voronezh State University, 1 Universitetskaya pl., Voronezh, 394018, Russian Federation
³Joint Stock Company «Concern «Sozvezdie»»,14 Plehanovskaya, Voronezh, 394018, Russian Federation

E-mail: vip@phys.vsu.ru

Abstract. The problem of estimating unknown signal parameters in the presence of noise is a classical problem in statistical radio engineering. There are many algorithms for solving this problem. Many of these algorithms are based on the use of well-known statistical optimality criteria in order to improve the estimation efficiency. Therefore, they acquire a high computational complexity. In this article, in order to reduce computational complexity, it is proposed to modernize the classical correlation algorithm based on the new Compressive Sensing theory, which has been actively developed in recent decades. The article presents a new simplified algorithm of signal parameters estimation, outlines operational speed benefits, and compares the efficiency of the proposed algorithm to the efficiency of the classical algorithm as well as some other algorithms. The proposed algorithm has also been compared to the classical one according to computational complexity. Practical implementation of the proposed algorithm requires much smaller number of arithmetical operations with some deterioration in unknown parameter estimation accuracy.

1. Introduction

The problem of estimating signal parameters in the presence of noise has to be solved to contribute to the development of radar-tracking, navigation or telecommunication systems, including decentralized tasks in wireless sensor networks (WSN) etc. In this field, many different schemes for parameter estimation have been proposed, e.g. [1–13]. The major research trends are the following. First, it is the application of parameter estimation theory to different signal models, such as multiharmonic [6], polynomial phase signals [7], chirp-signals [8] etc. Another target of research is the influence of various (apart from Gaussian) disturbances, such as alpha-stable [9], impulsive [10] etc on parameter estimation efficiency. New algorithms are being developed, not only based on the principles of Maximum Likelihood theory, but also on Cluster analysis [11], Gabor [12], Radon [13] and other transforms. Recently, much attention has been given to the tasks of parameter estimation in distributed systems [14–16].
However, most of this methods may have heavy computational burden. On the other hand, it is well known from the related literature that it is possible to obtain a many fewer number of measurements without any loss of information provided that Compressive Sensing (CS) theory is used [17–19]. While most of the works in CS are focused on the recovery of the observed signal from its measurements, in the past few years there has been growing interest in the application of CS concepts to parameter estimation problems [20]. Nevertheless, the major focus is given to estimation of the following two parameters – time-delay and harmonics frequencies [21, 22]. For this purpose, algorithms normally classified as subspace methods are applied and further modified. All the named approaches have rather tenuous relationship to classical parameter estimation algorithms, which are mostly based on likelihood ratio [1–5]. Moreover, the performance of most parameter estimation methods is known to differ distinctly from Cramer-Rao bound at low signal to noise ratio (SNR). This phenomenon is called a subspace swap. In this article, we propose a new algorithm that has better performance than CS’s in such conditions.

The main contributions with respect to existing literature on parameter estimation problem can be highlighted as follows. In CS theory, the problem of parameter estimation is often formulated as a problem of estimating the positions and amplitudes of nonzero components of a sparse signal from compressed data (if the original signal itself is sparse) or as the problem of estimating numbers and amplitudes of spectral coefficients (if the signal is sparse in the frequency domain). There are many algorithms for solving this problem [19]. At the same time, the problem of estimating the signal parameters can be solved using the classical Bayesian approach or the asymptotically optimal maximum likelihood algorithm [1–5]. The present paper describes an attempt to combine the classical (based on likelihood ratio) approach to parameter estimation with CS theory. Indeed, we consider likelihood ratio as a function of an unknown parameter within a wide prior interval of its possible values. Meanwhile, provided that SNR value is not too small, the vector constructed from likelihood ratio references, can be considered as a compressed vector. Consequently, the CS theory should be applied to likelihood ratio, but not to the signal itself. Taking into account that measurement matrix [17–19] can be formed in advance (before the useful signal arrival), the structure of the proposed algorithm turns out to have little difference from the classical correlation algorithm. Meanwhile, its computational complexity becomes significantly lower than the one of a classical algorithm.

We formulate the parameter estimation problem as follows. Assume that over the interval \([0;T]\) the observable data \(x(t)\) represents an additive mixture of a hindrance \(n(t)\) and a signal \(s(t, \Theta^{(0)})\), where \(\Theta^{(0)}\) represents the true values of the signal parameter vector. Assume that these parameters are unknown (vector \(\Theta\)); only a priori domains of their possible values are known: \(\Theta_r \in [\Theta_{\text{min}}; \Theta_{\text{max}}]\), \(r = 1, \ldots, R\), where \(R\) is the number of unknown parameters. It is required to estimate, or measure, the values of the unknown parameters (vector \(\Theta\)) of a signal according to observable data \(x(t)\). Then we suggest that purely discrete processing should be done. Thus, the observable data are expressed via vector \(\mathbf{x} \in \mathbb{R}^N\), and the number of this vector time samples \(N\) are picked out according to Whittaker–Kotelnikov–Shannon theorem [19] to minimize the loss from sampling. To solve the formulated problem, a considerable number of various algorithms have been developed [1–5]. They depend on the a priori information and the chosen criterion of estimation optimality. In case of Gaussian hindrance statistics \(\mathbf{n} \in \mathbb{R}^N\), the most productive is the approach based on estimating the signal parameters according to the maximum likelihood ratio [1–5]. For example, if \(\mathbf{n}\) is the vector containing Gaussian zero mean independent random variables with known variance \(D\), the log-likelihood ratio will be represented similar to [1] as

\[
M(\Theta) = D^{-1} \sum_{i=1}^{N} x_i s_i(\Theta) - (2D)^{-1} \sum_{i=1}^{N} s_i^2(\Theta),
\]

where \(x_i = x(t_i), s_i(\Theta) = s(t_i, \Theta)\).
Thus, the estimates of unknown parameters \( \hat{\theta}_r \) \((r=1,\ldots,R)\) are defined by the position of the absolute (greatest) maximum of solving statistics (1). However, an analytical search of the maximum of function for many variables is almost impossible. Therefore, this search has to be carried out by numerical methods, which requires sampling (1) for all unknown parameters \( \theta_r \) \((r=1,\ldots,R)\). For this purpose the a priori domain \([\Theta_{r_{\min}};\Theta_{r_{\max}}]\) of possible values for each parameter is split into \(N_{\theta_r}-1\) subintervals of the same length \(\Delta \theta_r\). Then the values of the log-likelihood ratio (1) for all possible combinations of unknown parameters \( \theta_r \in [\Theta_{r_{\min}};\Theta_{r_{\max}}], \ r = 1,\ldots,R, \) each of which is sampled with a step \(\Delta \theta_r \) \((r=1,\ldots,R)\), are calculated. As a result, the array of solving statistics (1) values of size \(N_x = N_{\theta_1}N_{\theta_2}\ldots N_{\theta_R}\) is obtained, among which the maximal one is to be found. The unknown parameters estimates are defined by the indices of each parameter at which the absolute maximum value is reached. The value of \(N_x\) can reach huge values. That occurs, for instance, at the initial stage of signal search, when a priori intervals \(\Theta_{r_{\max}} - \Theta_{r_{\min}}\) are large. Consequently, the number of arithmetical operations required to define all \(N_x\) values of the solving statistics can increase so much that either expensive high-speed processors will be required, or the decision on parameters estimates will be taken with some delay. Moreover, the less the processor speed is, the greater the delay will be.

Thus, there is a need for estimation algorithms that would not require so numerous arithmetical operations as classical algorithms described above. The main objective of the present article is to describe the development strategy of such algorithms.

2. Materials and methods

2.1. Main principles of Compressive Sensing theory

Compressive Sensing is a new concept in signal processing and information theory, where a small number of non-adaptive linear combinations of the signal are measured. The CS theory considers signals, which have sparse or compressible structure. Assume that \( x \) is a discrete time signal, which can be viewed as an \( N \times 1 \) column vector in \( \mathbb{R}^N \). Assume that this vector is directly (or in some basis) sparse, i.e. possesses only nonzero elements \( k \) \((k \ll N)\). Typically, real-world signals are not exactly sparse in any orthogonal basis. Instead, they are compressible. A signal is said to be compressible if the most of its elements are sufficiently small, and only a few has the large magnitudes.

Then, the measurement matrix should be selected. In CS theory it is proved that as such a matrix one can choose a matrix \( \Phi \) of size \( M \times N \) \((M \ll N)\) with independent and identically distributed normal random variables.

Since \( M \ll N \), we have an under-determined system of linear equations with an infinite number of solutions. Therefore, our problem is ill-posed. The main approach taken in CS is to minimize the \( l_1 \)-norm

\[
\hat{x} = \arg \min_x \|x\|_1 \quad s.t. \quad y = \Phi x'.
\]

This convex optimization program is often known as Basis Pursuit [19]. This equation can be solved via linear programming algorithms [18, 19] and several other algorithms [19]. Similarly, the problem is solved if there are observations corrupted by noise [17, 19].

So, if the discrete signal is sparse (or compressible) in some basis, it can be represented by a reduced number of samples compared to what the Whittaker–Kotelnikov–Shannon theorem suggests.

2.2. Application of Compressive Sensing theory for parameter estimation

Let us consider how this methodology can be applied to solve the problem of parameter estimation. We will limit ourselves to the situation when only one parameter \( \theta \) is unknown, i.e. \( R=1 \). The
multidimensional case generalization ($R > 1$) is obvious. We will present solving statistics (1) in the form of a vector of size $N_\theta \times 1$:

$$M = A - B,$$  \hspace{1cm}  \text{(3)}

where $A = Wx$; $W$ is a matrix of size $N_\theta \times N_i$ with elements $\|s(t, \theta)\|$, $i = 1, ..., N_\theta$, $j = 1, ..., N_i$; $B$ is a vector of size $N_\theta \times 1$ with elements $\sum_{j \neq i} s^2(t, \theta) / 2$, $i = 1, ..., N_\theta$. In the resulting formulas $N_i$ is the number of observable data samples in the time domain; $N_\theta$ is the number of samples of the unknown parameter $\theta$ within the a priori interval $[\Theta_{\min}, \Theta_{\max}]$. The estimate $\hat{\theta}^{(CL)}$ of the unknown parameter will be thus defined by the number of the maximum sample in the vector $M$.

As mentioned earlier, in many cases, for example, in WSN nodes, it is impractical to use complex algorithms that require a large number of arithmetic operations. In this regard, let us estimate the number of operations of scalar multiplication and addition necessary for vector $M$ (3) construction. For this purpose, we will consider that vector $B$ can be calculated beforehand (before vector $x$ is obtained). Therefore, the most complicated task will be to create vector $A$ elements in (3). Indeed, in this case $N_\theta N_i$ multiplication operations and $N_\theta (N_i - 1)$ addition operations are required. As stated previously, the number of samples, both time and unknown parameter $\theta$ ones, and hence the number of required arithmetical operations, can appear too large as a result. We will consider the possibility to reduce the number of arithmetical operations used without essential efficiency deterioration in unknown parameter estimation. For that purpose, we will apply the approach based on CS methodology.

As abovementioned, vector $B$ in (3) can be calculated beforehand. Besides, if parameter $\theta$ is non-power [4], this vector does not influence on the estimation of this parameter. Therefore, further attention has to be given to vector $A$ construction. So, let us apply the solving statistics data, having presented them as $M = A = Wx$. Then we will represent that as

$$M = S + N,$$  \hspace{1cm}  \text{(4)}

where $S = WS^{(0)}$ is a signal vector, $N = Wn$ is a noise vector, $S^{(0)}$ is a vector of size $N_i \times 1$ with $\|s(t, \theta^{(0)})\|$, $j = 1, ..., N_i$ elements, and $\theta^{(0)}$ is the true value of the unknown parameter $\theta$.

The signal vector $S$ (mean of the solving statistics $M$) is a vector of $N_\theta$ elements that contains a number of elements $P$ of essential values, closed by the true value of parameter $\theta^{(0)}$; the other elements of this vector accept the values close to zero. Hence, $S$ is a compressible vector, and vector $M$ is the sum of compressible and noise vectors that corresponds to standard conditions of the compressive sensing theory. According to this methodology, we generate a measurement matrix $\Phi$ of size $M_\theta \times N_\theta$ where $M_\theta < N_\theta$, that satisfies the necessary conditions [17, 19]. Then, we construct a vector

$$L = \Phi M = \Phi S + \Phi N = \Phi S + N',$$  \hspace{1cm}  \text{(5)}

where $N' = \Phi N$ is a noise vector of size $M_\theta \times 1$.

The comparison of equation (5) and equation (2) allows realizing that we come to the main problem in the compressive sensing theory. Then, having created the observation vector $L$ of size $M_\theta \times 1$, we recover vector $\hat{S}$ of size $N_\theta \times 1$ via standard procedures of CS methodology [17–19], with the position of the maximum element in this vector defining the unknown parameter estimate $\hat{\theta}^{(CS)}$. 
Let us now outline the benefits of such approach in comparison with the classical one (based on maximum $M$). For that purpose, we will define the number of arithmetical operations necessary for vector $L$ construction via equation (5). To do that we will represent equation (5) as

$$L = \Phi M = \Phi W x = \Psi x,$$

where $\Psi = \Phi W$ is a matrix of size $M_\theta \times N_t$, which can be calculated beforehand (before vector $x$ is formed). Hence, to create vector $L$, $M_\theta N_t$ multiplication operations and $M_\theta (N_t - 1)$ addition operations are required. If $M_\theta << N_\theta$, the number of mathematical operations required to construct a vector $L$ is much fewer than the number of operations used for the classical approach (by vector $M$ formation (3)).

Despite qualitative reasoning presented herein, the solution of the problem of parameter estimation based on CS methodology is not fully satisfactory. First, except for vector $L$ (6) calculation, for which a reduced, in comparison with the classical algorithm, number of arithmetical operations are required, this method also suggests solving type (2) optimizing problem, which still requires both hardware and time expenditures. Secondly, as calculations show, with the increase in compression ratio (parameter $M_\theta / N_\theta$ reduction) and with SNR growth $\hat{\theta}^{(CS)}$ estimate characteristics can significantly concede to the characteristics of the classical estimate $\hat{\theta}^{(CL)}$. For instance, at $M_\theta / N_\theta = 0.5$ and SNR $z = 5$, the $\hat{\theta}^{(CS)}$ estimate variance will be 10 times greater than the $\hat{\theta}^{(CL)}$ estimate variance and this loss in estimation accuracy will increase up to 100 at $z = 7$ already.

2.3. Signal parameter estimation algorithm with reduced number of samples

We offer another approach to the problem of discrete signal parameter estimate $\theta$. As stated previously, the signal vector $S$ in equation (5) contains a limited number $P << N_\theta$ of the nonzero samples concentrated around the true value of parameter $\theta^{(0)}$. Thus, with the reconstruction of nonzero samples of the vector $S$ (type (2) optimization problem solution) the reconstruction accuracy deteriorates with the increase of $P$. Instead of restoring each element of the vector $S$, we offer to restore the summed values of the elements in some ranges. In other words, if to sum up the first $P$ elements, then the second $P$ elements, etc. in the vector $S$ and to construct the vector $g$ with the values obtained, this vector will contain only one or two nonzero elements instead of $P$ elements of great values otherwise. This procedure can be mathematically described as follows. Let us create a matrix $D$ of size $L_\theta \times N_\theta$, where each row contains $P$ units, and the other elements are equal to zero. Here, $L_\theta = [N_\theta / P]$, where $[x]$ rounds the element $x$ to the nearest integer. In the $i$-th row $(i = 1, \ldots, L_\theta)$ the units begin at $P(i - 1) + 1$ position and end at $P \cdot i$. For example, if $P = 5$, the matrix $D$ looks as follows:

$$D = \begin{pmatrix}
1111100000000000 \ldots 00000 \\
00000111110000 \ldots 00000 \\
\ldots \ldots \ldots \ldots \\
00000000000000 \ldots 11111
\end{pmatrix}.$$

Then we will create a vector $g$ of size $L_\theta \times 1$ as $g = DS$. Hence, $S = D^T g$, where the «$^T$» sign designates the pseudo-inverse operation of a matrix [19], i.e. $D^T = D^T (D D^T)^{-1}$, and «$\dagger$» stands for the transpose operation. As a result, the vector $L$ in equation (5) will be represented as

$$L = \Phi D^\dagger g + N' = F g + N',$$
where \( F = \Phi D^\dagger \) is the matrix of size \( M_\theta \times L_\theta \).

Assume that \( L_\theta \ll M_\theta \). Hence, equation (8) is possible to be viewed as the overdetermined linear system of equations with the number of unknown values is less than the number of the equations. A similar system of equations can be solved via

\[
g^{(D)} = F^\dagger L_\theta.
\]

Then the parameter \( \theta \) is estimated via

\[
\theta^{(D)} = \arg \max_{D} g^{(D)} / L_\theta.
\]  

Let us estimate the number of arithmetical operations required to find the estimate (10). For this purpose, represent equation (9) as \( g^{(D)} = (\Phi D^\dagger)^\dagger \Phi W x = T x \), where \( T = (\Phi D^\dagger)^\dagger \Phi W \) is a matrix of size \( L_\theta \times N_t \), the elements of which can be calculated beforehand. Hence, the construction of the vector \( g \) requires \( L_\theta N_t \) operations of multiplication and \( L_\theta (N_t - 1) \) addition operations. To complete the estimation, one more multiplication by number \( 1/L_\theta \) is necessary.

2.4. Analysis of computational complexity of algorithms

Thus, we have three algorithms to estimate a parameter \( \theta \): classical correlation algorithm (algorithm CL) based on searching the number of vector \( M \) maximal element (4); standard CS theory-based algorithm (algorithm CS) based on searching the number of vector \( L \) maximal element (6); and the proposed algorithm (algorithm D).

Taking into account that \( N_\theta > M_\theta > L_\theta \), it can be easily noticed that the proposed algorithm is possible to implement in a simpler way than algorithm CL and algorithm CS.

We first analyze the computational complexity of algorithm D in terms of the number of multiplications. Then we compare the computational complexity of this method with the algorithm CL and algorithm CS with the parameters considered in this work (figure 1).

\[\begin{array}{c c c}
\text{CL} & \text{CS} & \text{D} \\
\text{~2 million} & \text{~1 million} & \text{~65 thousand} \\
0 & 2000000 & 2000000 \\
1000000 & 1500000 & 1500000 \\
2000000 & 2500000 & 2500000 \\
\end{array}\]

**Figure 1.** The diagram of the computational complexity of the algorithms (the number of multiplication operations).

The computational costs (the number of multiplications and additions \( N_\lambda \)) of the above three methods are listed in table 1. Taking into account that \( N_\theta \gg M_\theta \gg L_\theta \), it is evident that the computational complexity of the proposed method is much lower than that of the others.
Table 1. Computational complexity of algorithms.

| Algorithm | CL | CS | D |
|-----------|----|----|---|
| Computation complexity \( (N_a) \) | \( 2N_oN_i \) | \( 2M_oN_i \) | \( 2L_oN_i \) |

Then the computational cost of the presented method is studied and compared with the existing parameter estimation methods. The averaged CPU time of the estimation process over 100 simulations is used as a rough measure of the complexity of the estimation algorithm for different values of parameter \( M_o / N_o \). Simulations are performed in MATLAB 7.1 environment, using AMD Athlon (TM) 64 Processor 3000+, 1.81 GHz processor with 704 MB of memory, and 32 bit Microsoft Windows XP Professional operating system. The results are shown in table 2.

Table 2. Averaged CPU time of the estimation process for the considered algorithms (milliseconds).

| \( M_o / N_o \) | 0.1 | 0.3 | 0.8 |
|-----------------|-----|-----|-----|
| CL              | 15  | 15  | 15  |
| CS              | 33  | 110 | 410 |
| D               | 0.4 | 0.4 | 0.4 |

It is obviously seen that the classical and proposed algorithm estimation times do not depend on parameter \( M_o / N_o \), named compression ratio. The mean estimation time for CS algorithm sufficiently rose with the compression ratio increase. Based on the foregoing, algorithm CS was not considered further.

2.5. Efficiency of unknown parameter estimation via the classical algorithm

Now compare the efficiency of the parameter \( \theta \) estimation via classical in contrast with the proposed algorithm. We will describe this efficiency by such characteristics as unconditional bias and variance of the corresponding estimate. As an example, represent the signal as \( s(t, \theta) = A \exp\left(-\beta^2(t-\theta)^2\right) \), where the unknown parameter \( \theta \) is the time delay. It is theoretically possible to define only the bottom bound of the \( \hat{\theta}^{(CL)} \) estimate variance known as Cramer-Rao bound [4]. Hence, for the classical algorithm based on the statistics (1), the estimate \( \hat{\theta}^{(CL)} \) is asymptotically (with SNR growth) unbiased and its unconditional variance is at least \( V_0 \):

\[
V_0^{(CL)} \geq V_0 = \int \frac{p(\theta)}{-\left\{ \partial^2 M(\theta)/\partial \theta^2 \right\}} d\theta.
\]

Here \( p(\theta) \) is a priori probability density of the unknown parameter, assumed further uniform on the \([\Theta_{\text{min}}, \Theta_{\text{max}}]\) interval, with angular brackets indicating the average operator. It can be easily defined that the bottom bound for unconditional variance of a normalized parameter \( \tilde{\theta} = \theta / \Theta_{\text{max}} \) estimate looks like

\[
V_0 = \frac{\sqrt{2\pi}}{\xi^2(1-\gamma_1/(2\xi))\gamma_1 \gamma_2(\gamma_1+\gamma_2)\sum_{i=1}^{N_a} \left[ \gamma(i-1)-2\xi x \right]^2 \exp\left[-(\gamma(i-1)-2\xi x)^2/2 \right]} dx.
\]
Here $\xi = \beta \Theta_{\text{max}}$; $\gamma_1 = 0.1$ is the parameter that defines a sampling interval $\Delta \tilde{\theta}$ for the parameter $\tilde{\theta}$ (it is assumed that this interval is 10 times less than the correlation interval of noise function $M(\tilde{\theta}) - \langle M(\tilde{\theta}) \rangle$, that guarantees no more than 5% mean square error at $M(\tilde{\theta})$ construction); $\gamma_2 = 3$ is the parameter that defines the width of a signal function $\langle M(\tilde{\theta}) \rangle$; $N_t = 1 + (2\xi + \gamma_2) / \gamma_1$ is the number of time samples for the observation interval.

Note that the formula (12) is correct only at SNR $z \to \infty$, when abnormal errors can be neglected. As abnormal errors, such errors are understood with which the unknown parameter estimate falls outside the limits of the signal function $S$ width. Taking into account abnormal errors, unconditional variance of the parameter $\tilde{\theta}$ estimate will be represented as [4, 6] $V^{(CL)} = P_0 V_\theta + (1 - P_0)V_A$, where

$$P_0 \approx 1 - \frac{\xi}{2\pi\sqrt{3}} \exp(-z^2 / 4),$$

$$V_A = -\frac{1}{2} \left(1 + \frac{\gamma_2}{2\xi}\right)^2 + \frac{2}{3} \left(1 + \left(\frac{\gamma_2}{2\xi}\right)^2 + \frac{\gamma_2}{2\xi}\right),$$

and $V_\theta$ is defined via the formula (12).

3. Results and discussion

In figure 2 the continuous line represents the dependence of unconditional variance $V^{(CL)}$ for algorithm CL, and the dash-dot line represents the unconditional variance $V_\theta$ corresponding to the bottom of Cramer-Rao bound, from SNR $z$. The analysis of the figure allows concluding that with SNR values $z < 8$ the variance of estimate (with regard to abnormal errors) much exceeds the Cramer-Rao bound variance. Therefore, Cramer-Rao bound will not be further applied, and all calculations will be done with regard to abnormal errors.

It is rather difficult to estimate the efficiency of the algorithm D analytically. Therefore, its efficiency was tested by computer simulations in MATLAB. The simulations were organized as follows. Under conditions similar to the ones previously described, the vector (9) was created with regard to equation (8) and equation (6), then its maximum element was found and finally the estimation was defined according to equation (10). The simulation results are represented in figure 2 by the dashed line at $M_\theta / N_\theta = 0.5$.

![Figure 2. Unconditional variance of estimate versus SNR.](image-url)
As research shows, there is little change in variance of estimate $V^{(D)}$ with the change of parameter $M_\theta / N_\theta$ within $0.1 \rightarrow 1$ limits. Obviously, with SNR growth the algorithm D suggests poorer estimation accuracy compared to the classical algorithm. However, as it was noted before, the algorithm D essentially benefits in comparison with the classical algorithm based on solving statistics (1) by the number of arithmetical operations required. Indeed, under the simulation conditions, the number of time samples and parameter $\tilde{\theta}$ samples were equal to $N_t = 1972, N_\theta = 986$ accordingly. Hence, the number of multiplication operations necessary for statistics (1) should be about 2 million. At the same time, the application of algorithm (10) with $L_\theta = 33$, decreases the number of similar operations significantly, i.e. to about 65 thousand. As a result, the approximate number of multiplication operations is 30 times less (there is a similar decrease in the approximate number of addition operations as well). Note that the benefit in the number of operations will be ever more with the increase in the a priori interval of possible parameter $\tilde{\theta}$ values, i.e. with the increase in the parameter $\xi$ value.

The number of arithmetical operations for the classical algorithm based on the statistics (1), can be reduced by reducing the number of parameter $\tilde{\theta}$ samples $N_\theta$, i.e. by increasing the sample interval $\Delta \tilde{\theta}$. For that purpose, we will introduce the $K$ parameter, which takes the integer positive values and defines the decimation degree of statistics (1) for parameter $\tilde{\theta}$ (the sample interval thus increases in $K$ times). Let us first look into the dependence of the parameter $\tilde{\theta}$ estimation efficiency on parameter $K$. Figure 3 shows the unconditional variance of the parameter $\tilde{\theta}$ estimate versus decimation parameter $K$ value for classical algorithm.

![Figure 3. Unconditional variance of estimate versus decimation parameter $K$.](image)

As follows from the analysis of this figure, the parameter $\tilde{\theta}$ decimation leads to essential growth in the variance of estimate already at $K > 3 + 5$, regardless of SNR. However, with the growth of $K$ value, the number of arithmetical operations for such algorithm will be reduced in $K$ times correspondingly. Also it was shown that the estimation efficiency of proposed algorithm D weakly depends from parameter $\kappa = M_\theta / N_\theta$ if $\kappa > 0.8 + 1$.

Let us now compare the estimations efficiency for the algorithm D and for the decimation algorithm. In figure 4 the continuous lines show unconditional variance of estimate $\tilde{\theta}^{(D)}$ as a function of SNR $z$ for proposed algorithm D. Comparatively, the dashed lines in the same figure represent similar dependencies for algorithm CL with decimation.
The analysis of this figure shows that in case of decimation, the number of required arithmetical operations naturally decreases, however this considerably worsens the estimation efficiency.

For example, at $K = 2$ the decimation algorithm is less efficient than algorithm D at $z > 7$ and $\kappa > 0.8$, with a 15 times greater number of arithmetical operations with the first algorithm compared to the second algorithm.

If to perform the decimation with even greater degree (e.g. with $K = 5$) the algorithm D will start possessing the best noise immunity (smaller variance of estimate) at $z > 5$ and $\kappa > 0.1$. The decimation algorithm will still lose to algorithm D as for the number of arithmetical operations, the loss being approximately six times greater number of operations required.

If to decimate the parameter $\theta$ statistics (1) so that the number of arithmetical operations for both algorithms would become identical, the accuracy of the decimation algorithm will substantially concede to the algorithm D at $\kappa > 0.1$ and at any $z$ values already. Thus, the proposed algorithm D can be effectively applied in case of rigid restrictions on the operational speed of the processor.

4. Conclusion
It is shown that the proposed algorithm for estimating signal parameters significantly gain in speed in comparison with the classical algorithm. This gain increases with the value of the a priori interval of possible values of the estimated parameter. Reducing the required number of arithmetic operations in the classical algorithm due to decimation leads to a significant loss in the estimation efficiency compared to the proposed algorithm. The gain in speed of the proposed algorithm in comparison with the classical one will be even greater if not one but several parameters of the signal are estimated.

References
[1] Van Trees H L and Bell K L 2013 Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory. 2nd ed. (New York: Wiley) p 1176
[2] Van Trees H L 2001 Detection, Estimation, and Modulation Theory, Part III: Radar-Sonar Signal Processing and Gaussian Signals in Noise (New York: Wiley) p 626
[3] Barkat M 2005 Signal Detection and Estimation. 2nd ed. (Boston: Artech House) p 692
[4] Kulikov E I and Trifonov A P 1978 Estimation of the Parameters of Signals Against the Background of Interference (Moscow: Sovetskoe Radio) p 296
[5] Kay S M 1993 Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory (New Jersey: Prentice Hall) p 595
[6] Beltran-Carbajal F and Silva-Navarro G 2017 A fast parametric estimation approach of signals with multiple frequency harmonics Electric Power System Research 144 pp 157–62

[7] Ou G and Liu Z 2017 A modified CPF algorithm for estimating the parameters of multicomponent third-order polynomial phase signals Proc. of IEEE 17th International Conference on Communication Technology (ICCT) pp 338–42

[8] Al-Irkhis L A 2017 Fast CS algorithm for chirp signals parameters estimation using multi-resolution DCFT atom transformation Proc. of IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON) pp 31–6

[9] Yang G, Wang J, Zhang G, Shao O and Li S 2018 Joint estimation of timing and carrier phase offsets for MSK signals in alpha-stable noise IEEE Communications Letters 22 1 pp 89–92

[10] Alam M S, Kaddoum G and Agba B L 2018 Performance analysis of distributed wireless sensor networks for gaussian source estimation in the presence of impulsive noise IEEE Signal Processing Letters 25 6 pp 803–7

[11] Mo D and Duarte M F 2018 Compressive parameter estimation via K-median clustering Signal Processing 142 pp 36–52

[12] Lu Z, Yang C, Qin D and Luo Y 2016 Estimating the parameters of ultrasonic echo signal in the Gabor transform domain and its resolution analysis Signal Processing 120 pp 607–19

[13] Wang Z, Wang Y and Xu L 2017 Parameter estimation of hybrid linear frequency modulation-sinusoidal frequency modulation signal IEEE Signal Processing Letters 24 8 pp 1238–41

[14] Katenka N, Levina E and Michailidis G 2013 Tracking multiple targets using binary decisions from wireless sensor networks Journal of the American Statistical Association 108 502 pp 398–410

[15] Al-Jarrah M A, Al-Ababneh N K and Al-Ibrahim M M 2014 Cooperative OFDM for semi-distributed detection in wireless sensor networks AUE – International Journal of Electronics and Communications 68 10 pp 1022–9

[16] Althunibat S, Khalifeh A and Mesleh R 2018 A low-interference decision-gathering scheme for critical event detection in clustered wireless sensor network Physical Communication 26 pp 149–55

[17] Candès E J and Wakin M B 2008 An introduction to compressive sensing IEEE Signal Processing Magazine 25 2 pp 21–30

[18] Donoho D L 2006 Compressed sensing IEEE Transaction on Information Theory 52 4 pp 1289–306

[19] Foucart S and Rauhut H 2013 A Mathematical Introduction to Compressive Sensing (Springer) p 625

[20] Wu S, Tian J and Cui W 2015 A novel parameter estimation algorithm for DSSS signals based on compressed sensing Chinese Journal of Electronics 24 2 pp 434–8

[21] Chan Y T, Chan F, Rajan S and Lee B H 2015 Direct estimation of time difference of arrival from compressive sensing measurements Proc. of 3rd International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa) pp 273–6

[22] Huang X, Zhao Q, Jiang W, Zou L, Gui G and Wan Q 2015 Frequency estimation of cyclic spectrum carrier based on compressive sampling of BPSK signal Proc. of IET International Radar Conference pp 1–4