Research Article

On a Gradient-Based Algorithm for Sparse Signal Reconstruction in the Signal/Measurements Domain

Ljubiša Stanković and Miloš Daković

University of Montenegro, 81000 Podgorica, Montenegro

Correspondence should be addressed to Ljubiša Stanković; ljubisa@ac.me

Received 16 March 2016; Accepted 24 May 2016

Academic Editor: Cornel Ioana

Copyright © 2016 L. Stanković and M. Daković. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sparse signals can be recovered from a reduced set of samples by using compressive sensing algorithms. In common compressive sensing methods the signal is recovered in the sparsity domain. A method for the reconstruction of sparse signals which reconstructs the missing/unavailable samples/measurements is recently proposed. This method can be efficiently used in signal processing applications where a complete set of signal samples exists. The missing samples are considered as the minimization variables, while the available samples are fixed. Reconstruction of the unavailable signal samples/measurements is performed using a gradient-based algorithm in the time domain, with an adaptive step. Performance of this algorithm with respect to the step-size and convergence are analyzed and a criterion for the step-size adaptation is proposed in this paper. The step adaptation is based on the gradient direction angles. Illustrative examples and statistical study are presented. Computational efficiency of this algorithm is compared with other two commonly used gradient algorithms that reconstruct signal in the sparsity domain. Uniqueness of the recovered signal is checked using a recently introduced theorem. The algorithm application to the reconstruction of highly corrupted images is presented as well.

1. Introduction

A signal is sparse in a transformation domain if the number of nonzero coefficients is much lower than the number of signal samples. For linear signal transforms the signal samples can be considered as linear combinations (measurements) of the signal transform coefficients. Sparse signals can be fully recovered from a reduced set of samples/measurements. Compressive sensing theory is dealing with sparse signal reconstruction [1–22]. In common signal processing problems, for signals sparse in a transformation domain, signal samples can be considered as observations since they are linear combinations of the transformation domain coefficients. The goal of compressive sensing is to reduce the set of signal measurements while preserving complete information about the whole signal. In some applications unavailable signal samples result from the system physical constraints including unavailability to perform all signal measurements. Signals may also contain arbitrarily positioned samples that are heavily corrupted. In these cases it is better to remove such samples and consider them as unavailable in the signal recovery [6]. Compressive sensing theory based reconstruction algorithms may be also used when the unavailable signal samples are not a result of an intentional strategy to compress the data. Note that one of the initial compressive sensing theory successes in applications (computed tomography reconstruction) was not related to the intentional compressive sensing strategy but to the physical problem constraints, restricting the set and positions of available data.

Various algorithms for reconstruction of sparse signals from a reduced set of observations are introduced [8–17]. Two of the most important approaches are based either on the gradient of the sparsity measure [8] or on the orthogonal matching pursuit methods [16].

The topic of this paper is the reconstruction of signals with some arbitrary positioned unavailable samples. These samples may result from a compressive sensing strategy or from their unavailability due to various reasons. A method for the unavailable signal samples reconstruction, considering them as variables, has been proposed in [22]. In contrast to the other common methods that recover sparse signals in their sparsity domain this method reconstructs...
missing samples/measurements to make the set of samples/measurements complete. The available samples are fixed (invariant) during the reconstruction process. An analysis of the algorithm performance is presented in this paper. A relation between reconstructed signal bias and the algorithm step-size is derived and statistically confirmed. A new criterion for the algorithm step-size adaptation is proposed. The improved version of this algorithm is compared with two other common gradient-based algorithms proving its accuracy and computational efficiency. The discrete Fourier transform domain is used as a case study, although the algorithm application is not restricted to this transform [23]. The solution uniqueness is checked after signal reconstruction by using the recently proposed theorem [24].

The paper is organized as follows. After the definitions in the next section, the adaptive gradient algorithm is presented. In the comments to the algorithm the step-size and bias are considered and illustrated. A new criterion for the algorithm parameter adaptation is proposed in Section 3 as well. Reconstruction is illustrated on examples and uniqueness of the obtained solution is analyzed. Statistical performance are checked in Section 4. An illustration of the algorithm application to the image reconstruction is given as well.

2. Gradient-Based Reconstruction of Missing Samples

Discrete-time signal \( x(n) \) with \( N \) samples will be considered (with vector notation \( x = [x(0) \ x(1) \ \cdots \ x(N-1)]^T \)). Its transform coefficients will be denoted by \( X(k) = T[x(n)] \). If the number of nonzero coefficients \( X(k) \neq 0 \) for \( k \in \{k_1, k_2, \ldots, k_s\} \) is such that \( s \ll N \) then the signal \( x(n) \) is sparse in this transformation domain. Next assume that only \( M < N \) signal samples are available in the time domain at the discrete-time positions:

\[
n_j \in \mathbb{N}_A = \{n_1, n_2, \ldots, n_M\} \subset \mathbb{N} = \{0, 1, 2, \ldots, N-1\}.
\]

Under some conditions, studied within the compressive sensing theory [1–3], full signal reconstruction is possible based on the reduced set of available samples if the signal is sparse in a transformation domain. In the signal theory this problem can be formulated as a full recovery of the unavailable samples as well. In general, a reconstruction algorithm is based on the sparsity measure minimization. Nonzero transform coefficients’ counting is the simplest sparsity measure. Counting can be achieved by the mathematical form \( \|X\|_0 \), referred to as the “\( \ell_0 \)-norm” [1, 20]. The signal reconstruction problem statement is then

\[
\begin{align*}
\min & \quad \|X\|_0 \\
\text{subject to} & \quad y = AX.
\end{align*}
\]

(2)

Vector \( y = [x(n_1) \ x(n_2) \ \cdots \ x(n_M)]^T \) contains the available signal samples, while the transform vector \( X = [X(1) \ X(2) \ \cdots \ X(N)]^T \) elements are the unknown transform coefficients. Matrix \( A \) is the measurement matrix. In common signal transforms it corresponds to the inverse transformation matrix where the rows corresponding to the unavailable samples are omitted. If \( x = \Psi x \) then the matrix \( A \) is a submatrix of \( \Psi \) containing its rows corresponding to the discrete-time instants (1).

Although there are some approaches to reconstruct the signal using the \( \ell_0 \)-norm based formulation [25] in principle, this is an NP-hard combinatorial optimization problem. It is the reason why the closest convex \( \ell_1 \)-norm of the signal transform is used instead of the “\( \ell_0 \)-norm” in the minimization process:

\[
\begin{align*}
\min & \quad \|X\|_1 \\
\text{subject to} & \quad y = AX,
\end{align*}
\]

where \( \|X\|_1 = \sum_{k=0}^{N-1} |X(k)| \). Under the conditions defined using the restricted isometry property (RIP) [3, 4], minimization of the \( \ell_1 \)-norm can produce the same result as (2) [1, 22, 26]. Common gradient-based algorithms reformulate the problem defined by (3) into a form suitable for application of standard minimization tools. Minimization is done in the sparsity domain. One such a method is briefly reviewed in Appendix.

2.1. Algorithm. A gradient-based algorithm that minimizes the sparsity measure by varying the missing sample values is presented next. In this gradient-based minimization approach the missing samples are considered as variables [22]. The missing sample values are varied in an iterative way to produce lower sparsity measure values and finally to approach the minimum of the convex \( \ell_1 \)-norm based sparsity measure (3) with an acceptable accuracy. Note that if the reconstruction conditions are met for the \( \ell_1 \)-norm [3] then the \( \ell_1 \)-norm minimum position will correspond to the true values of the missing values.

The signal

\[
\begin{align*}
x_a^{(0)}(n) &= \begin{cases} 0 & \text{for missing samples, } n \in \mathbb{N}_Q \\ x(n) & \text{for available samples, } n \in \mathbb{N}_A \end{cases} \\
\end{align*}
\]

(4)

will be used as the initial estimate. Note that this signal would follow as a result of the \( \ell_1 \)-norm based minimization of the signal transform. The set of missing sample positions is \( \mathbb{N}_Q \). Values of the available samples are fixed and will be considered as constants. The missing sample values are varied through iterations. If we denote by \( x_a^{(m)} \) the vector of signal values reconstructed after \( m \) iterations, then the goal is to achieve

\[
\begin{align*}
\min & \quad \|X_a\|_1 \\
\text{subject to} & \quad x_a^{(m)}(n) = x(n) \quad \text{for } n \in \mathbb{N}_A,
\end{align*}
\]

(5)

where the components of \( X_a \) are \( X_a(k) = T[x_a^{(m)}(n)] \). In order to find the position of function \( z = \|X_a\|_1 \) minimum, the relation for iterative missing samples (variables) calculation is defined by using the gradient-based descend on the sparsity measure \( \|X_a\|_1 \) as

\[
\begin{align*}
y_c^{(m+1)} = y_c^{(m)} + \alpha \frac{\partial \|X_a\|_1}{\partial y_c}
|_{y_c=y_c^{(m)}}.
\end{align*}
\]

(6)
Vector \( y^{(m)} \) contains the minimization variables (missing signal sample values) in the \( m \)th iteration. The gradient vector coordinates \( g(n_i) = \partial \|X_n^i\| / (\partial y, N) \) can be estimated, in the \( m \)th iteration, by using the finite differences of the sparsity measure. They are calculated for each missing sample at instants \( n_i \in \mathbb{N}_Q \):

\[
g(n_i) = \frac{\|X_n^i\| - \|X_n^{i-1}\|}{2\Delta N}, \tag{7}
\]

where

\[
X_n^+(k) = T \{ x_n^+(n) \},
\]

\[
X_n^-(k) = T \{ x_n^-(n) \},
\]

\[
x_n^+(n) = X_n^{(m)}(n) + \Delta \delta(n - n_i),
\]

\[
x_n^-(n) = X_n^{(m)}(n) - \Delta \delta(n - n_i). \tag{9}
\]

Again for \( n_i \in \mathbb{N}_A \), the available signal values are not changed, \( g(n_i) = 0 \). All gradient vector \( g^{(m)} \) values, including the gradient zero and nonzero coordinates, in the \( m \)th iteration are denoted by \( g^{(m)}(n) \).

Through statistical study \cite{22} it has been concluded that an appropriate step parameter value in (6) is related to the finite difference step as \( \alpha = 2\Delta \).

\section{3. Comments on the Algorithm}

(i) The algorithm inputs are the signal length \( N \), the set of available samples \( \mathbb{N}_A \), and the available signal values \( x(n_i), n_i \in \mathbb{N}_A \).

(ii) For the DFT instead of using signals defined by (9) and their transforms for each \( n_i \in \mathbb{N}_Q \) it is possible to calculate

\[
|X_n^+(k)| = |X_n^{(m)}(k) + \Delta D_{n_i}(k)|,
\]

\[
|X_n^-(k)| = |X_n^{(m)}(k) - \Delta D_{n_i}(k)| \tag{10}
\]

with \( X_n^{(m)}(k) = \text{DFT}[x_n^{(m)}(n)] \) and \( D_{n_i}(k) = \text{DFT}[\delta(n - n_i)] = \exp(-j2\pi m k/N) \). Note that \( D_{n_i}(k) \) values are not dependent on the signal and the iteration number \( m \). They can be calculated only once. Similar simplification can be done for any linear signal transform \( T[.]. \)

\subsection{3.1. Adaptive Step}

The influence of step-size \( \Delta \) to the solution (minimum position) precision is analyzed next. Assume that we have obtained the exact solution for the missing samples. The change of sparsity measure is tested by the change of signal sample \( x(n_i) \) for \( \pm \Delta \). For a signal of sparsity \( s \) in the DFT domain, whose form is \( x(n) = \sum_{m=1}^{N} A_m e^{j2\pi m k/N} \), the signals with the steps used for the gradient estimate calculation are \( x_n^+(n) = x(n) + \Delta \delta(n - n_i) \) and \( x_n^-(n) = x(n) - \Delta \delta(n - n_i) \). Their sparsity measures are

\[
\|X_n^+(k)\| = \sum_{i=1}^{s} |A_i + \Delta e^{-j2\pi m k/N}| + (N - s) \Delta \tag{11}
\]

\[
\|X_n^-(k)\| = \sum_{i=1}^{s} |A_i - \Delta e^{-j2\pi m k/N}| + (N - s) \Delta.
\]

In the worst case, amplitudes \( A_i \) are in phase with \( e^{-j2\pi m k/N} \) and \( \Delta \leq |A_i| \) when

\[
\|X_n^+(k)\| = \sum_{i=1}^{s} |A_i| + s\Delta + (N - s) \Delta = \|X_n^+(k)\| = (N - s) \Delta
\]

\[
\|X_n^-(k)\| = \sum_{i=1}^{s} |A_i| - s\Delta + (N - s) \Delta
\]

\[
= \|X_n^+(k)\| + (N - 2s) \Delta. \tag{12}
\]

It means that \( g(n_i) = \|X_n^+(k)\| - \|X_n^-(k)\|/(2N\Delta) = s/N \neq 0 \). This is an important fact leading to the conclusion that the stationary state of the iterative algorithm is biased. The algorithm moves the solution \( x(n_i) \) to a biased value \( x(n_i) - b \) in order to produce \( g(n_i) = 0 \) in the algorithm stationary point. For the zero value of the gradient we have \( x_n^+(n_i) = x(n_i) - b + \Delta \) and \( x_n^-(n_i) = x(n_i) - b - \Delta \). By replacing \( \Delta \) with \( b \) and \( \Delta \) in \( \|X_n^+(k)\| \) and \( \|X_n^-(k)\| \) the worst case sparsity measures are \( \|X_n^+(k)\| = \|X_n^+(k)\| = (N - b) \Delta \) and \( \|X_n^-(k)\| = \|X_n^+(k)\| + (N - 2s) \Delta \). In this case the bias \( b \) value can be obtained from \( \|X_n^+(k)\| = \|X_n^+(k)\| \)

\[
N(\Delta - b) = (N - 2s)(\Delta + b)
\]

\[
b = \frac{s}{N - s} \Delta = \frac{s \Delta}{N} \quad \text{for} \ s \ll N. \tag{13}
\]

The bias limit is proportional to the step-size \( \Delta \). Obviously, the bias limit can be reduced by using small values of the step \( \Delta \). Calculation with a very small step \( \Delta \) would be inefficient and time consuming with an extremely large number of iterations. An efficient way to use the gradient-based algorithm is in adapting its step-size. In the initial iteration the step-size of an order of signal amplitude is used:

\[
\Delta = \max_{n \in \mathbb{N}_A} |x(n)|. \tag{14}
\]

When the algorithm reaches a stationary point with this step \( \Delta \) then the bias will dominate and the mean squared error will be almost constant. The algorithm can improve the solution precision by reducing the step-size.

This simple relation between the bias and step-size has been confirmed through statistical analysis on more complex cases, with a large number of missing samples.

\subsection{3.2. Adaptation Criterion}

The gradient-based algorithm convergence improvement by changing steps is analyzed in detail in \cite{22}. A criterion that can efficiently detect the event that the algorithm has reached the stationary point with
regard to the mean square error (the vicinity of the sparsity measure minimum defined by the bias) described in the previous subsection is proposed next. The presented criterion is based on the direction change of the gradient vector. When the vicinity of the minimum sparsity measure point is reached within the region defined by the bias value, then the gradient estimate oscillates and changes its direction for almost 180 degrees. The angle between two successive gradient vector estimations $g^{(m-1)}$ and $g^{(m)}$ denoted as $\beta_m$ can be calculated as
\[
\beta_m = \arccos \frac{\sum_{n=0}^{N-1} g^{(m-1)}(n) g^{(m)}(n)}{\sqrt{\sum_{n=0}^{N-1} (g^{(m-1)}(n))^2} \sqrt{\sum_{n=0}^{N-1} (g^{(m)}(n))^2}}.
\] (15)

If this angle $\beta_m$ is, for example, above 170° then we can conclude that the signal missing samples (variables) reached oscillatory nature around the minimal sparsity measure position. When these values of the angles are detected the algorithm step $\Delta$ is reduced, for example, for $\Delta/10 \rightarrow \Delta$ or $\Delta/\sqrt{10} \rightarrow \Delta$. The iterations are continued starting from the reconstructed signal values reached in the previous step.

When the minimum of the sparsity measure is reached with a sufficiently small $\Delta$, the value of $\Delta$ can also be used as an indicator of the solution precision. For example, if the signal amplitudes are of order 1 and $s/N = 0.1$ taking $\Delta = 1$ in the first iteration will produce the solution with a precision better than 20 [dB]. Then if the step $\Delta$ is reduced to $\Delta = 0.1$ a precision better than 40 [dB] will be obtained and so on. Value of $\Delta$ can be used as the algorithm stopping criterion.

### 3.3. Stopping Criterion

The precision of the result in iterative algorithms is commonly estimated based on the change of the result values in the last iterations. Therefore an average of changes in a large number of variables is a good estimate of the achieved precision. The value of
\[
T_r = 10 \log_{10} \frac{\sum_{n \in N_0} |x_a^{(p)}(n) - x_a^{(m)}(n)|^2}{\sum_{n \in N_0} |x_a^{(m)}(n)|^2}
\] (16)
can be used as a rough estimate of the reconstruction error to signal ratio. In this relation $x_a^{(p)}(n)$ is the signal reconstructed before $\Delta$ reduction and $x_a^{(m)}(n)$ is the reconstructed signal after iterations done with the reduced step. Value of $T_r$ can be used as the algorithm stopping criterion. If the value of $T_r$ is above the required precision $T_{r \text{max}}$ (e.g., if $T_r > -100$ dB), the calculation procedure should be continued with reduced values of step $\Delta$.

### 3.4. Large Step Influence

A possible divergence of a gradient-based algorithm is related to large steps $\alpha = 2\Delta$. We can write
\[
|X_a^{(p)}(k)| - |X_a^{(m)}(k)|
= |X_a^{(m)}(k) + \Delta D_{n_i}(k)| - |X_a^{(m)}(k) - \Delta D_{n_i}(k)|
= \Delta |D_{n_i}(k)| \left( \left| 1 + \frac{X_a^{(m)}(k)}{\Delta D_{n_i}(k)} \right| - \left| 1 - \frac{X_a^{(m)}(k)}{\Delta D_{n_i}(k)} \right| \right).
\] (17)

In general for a complex number $a = X_a^{(m)}(k)/(\Delta D_{n_i}(k))$, with $|a| < 1$ and a large $\Delta$, from the problem geometry we can show that the following bounds hold $0 \leq |1 + a| - |1 - a| \leq 2|a|$. Therefore,
\[
0 \leq |X_a^{(p)}(k)| - |X_a^{(m)}(k)| \leq 2 \left| X_a^{(m)}(k) \right|.
\] (18)

Lower limit 0 is obtained if $a$ is imaginary-valued, while the upper limit $2|X_a^{(m)}(k)|$ follows if $a$ is real-valued. If $\Delta$ is large the missing signal values will be adapted for a value independent on $\Delta$. The missing sample values will oscillate within the range of the original signal values until the step $\Delta$ is reduced in iterations. The variable values will be arbitrary changed within the signal amplitude order as far as $\Delta$ is too large. It will not influence further convergence of the algorithm, when the step $\Delta$ assumes appropriately small values.

A pseudocode of this algorithm is presented in Algorithm 1.

**Illustrative Example 1.** The effects analyzed within the comments on the algorithm will be illustrated on a simple signal:
\[
x(n) = 3 \cos \left( 2\pi \frac{n}{N} - \frac{\pi}{8} \right) - 0.2 \sin \left( 4\pi \frac{n}{N} + \frac{\pi}{2} \right)
\] (19)

with $N = 8$ and sparsity $s = 4$. Assume that the samples at positions $n \in N_{\Omega} = \{1, 6\}$ are missing. Reconstruction of the signal is done in 60 iterations. The initial value of the step $\Delta = 1$ and the initial values of missing samples $x(1) = 0$ and $x(6) = 0$ are used. The missing sample values in the first 20 iterations are given in Figure 1(a) by dots (connected by a line). After 6 iterations the algorithm does not significantly change the missing sample values with given step-size $\Delta = 1$ (lower subplot within the figure presents changes of variables in a highly zoomed scale). When the stationary point vicinity is reached with $\Delta = 1$ the gradient estimate is almost zero-valued. Its direction change for almost 180 degrees. The measure values are on the contour with approximately the same value (circles on contour plot). The algorithm resumes its fast approach toward the exact missing sample values after the algorithm step is reduced to $\Delta = 0.1$ in the 20th iteration. When a new stationary state is reached change of $\Delta$ to $\Delta = 0.01$ is done and the approach to the correct missing sample values is continued.

The bias upper limit in the stationary state for step $\Delta = 1$, according to (13), is $(s/(N - s))\Delta = 1$. After each reduction of the step-size from $\Delta$ to $\Delta/10$ the bias caused MSE will be reduced for almost 20 [dB]. The result for the signal reconstruction and the obtained MSE for the calculated missing samples $x(1)$ and $x(6)$ are presented in Figure 1.

**Illustrative Example 2.** The reconstruction process is repeated using the signal
\[
x(n) = 3 \sin \left( 20\pi \frac{n}{N} \right) + \cos \left( 60\pi \frac{n}{N} \right) + 0.7 \sin \left( 46\pi \frac{n}{N} \right)
\] (20)
Require:
(i) Set of missing/omitted sample positions \( N_\overline{Q} \)
(ii) Available samples \( x(n) \), \( n \notin N_\overline{Q} \)

1. Set \( x^{(0)}_a(n) \leftarrow x(n) \) \( \triangleright \) for \( n \notin N_\overline{Q} \)
2. Set \( x^{(0)}_Q(n) \leftarrow 0 \) \( \triangleright \) for \( n \in N_\overline{Q} \)
3. Set \( m \leftarrow 0 \)
4. Set \( \Delta \leftarrow \max \{ |x^{(0)}_a(n)| \} \)

repeat
6. Set \( x^{(m)}_a(n) = x^{(m)}(n) \)
7. repeat
8. \( m \leftarrow m + 1 \)
9. for \( n_i \leftarrow 0 \) to \( N - 1 \) do
10. if \( n_i \notin N_\overline{Q} \) then
11. \( X^{(m)}_a(k) \leftarrow \text{DFT}\{x^{(m)}_a(n) + \Delta \delta(n - n_i)\} \)
12. \( X^{(m)}_Q(k) \leftarrow \text{DFT}\{x^{(m)}_Q(n) - \Delta \delta(n - n_i)\} \)
13. \( g^{(m)}(n_i) \leftarrow (1/N) \sum_{k=0}^{N-1} |X^{(m)}_a(k)| - |X^{(m)}_Q(k)| \)
14. else
15. \( g^{(m)}(n_i) \leftarrow 0 \)
16. end if
17. \( x^{(m+1)}_a(n_i) \leftarrow x^{(m)}_a(n_i) - g^{(m)}(n_i) \)
18. end for
19. \( \beta_m = \arccos(\sum_{m=0}^{N-1} g^{(m)}(n)g^{(m)}(n)/\sqrt{\sum_{m=0}^{N-1} (g^{(m)}(n))}^2) \)
20. until \( \beta_m < 170^\circ \)
21. \( \Delta \leftarrow \Delta/\sqrt{10} \)
22. \( T_n = 10 \log_{10}(\sum_{n \in N_\overline{Q}} |x^{(m)}_a(n) - x^{(m)}_Q(n)|^2 / \sum_{n \in N_\overline{Q}} |x^{(m)}_a(n)|^2) \)
23. until \( T_n < T_n^{\max} \)
24. return \( x^{(m)}_a(n) \)

Output:
(i) Reconstructed signal \( x_a(n) = x^{(m)}_a(n) \)

Algorithm 1: Reconstruction.

and \( N = 32 \). Its sparsity is \( s = 6 \). Missing sample positions are at \( n \notin N_\overline{Q} =\{0,1,2,5,7,8,11,15,18,27,28,30\} \). The reconstruction results for this signal are presented in Figure 2. After the signal is reconstructed the problem is to establish if this reconstruction is unique in the sense that there is no other signal (missing sample values) with the same or lower sparsity. A very simple check of the solution uniqueness, using the missing sample positions and the positions of the reconstructed signal nonzero components in the sparsity domain, is given in [24] by the following.

Theorem 1. Consider a signal \( x(n) \) that is sparse in the DFT domain with unknown sparsity. Assume that the signal length is \( N = 2^h \) samples and that \( Q \) samples are missing at the instants \( q_m \in N_\overline{Q} \). 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 \( k_0 = \{k_0, k_02, \ldots, k_0s\} \). Reconstruction result is unique if the inequality

\[ s < N - \max_{h=0,1,\ldots,r-1} \left\{ 2^h (Q_{2^h} - 1) - s + 2S_{2^h} \right\} \]  

holds. Integers \( Q_{2^h} \) and \( S_{2^h} \) are calculated as

\[ Q_{2^h} = \max_{b=0,1,\ldots,2^h-1} \left\{ \text{card} \{ q : q \in N_\overline{Q}, \ mod \ (q, 2^h) \} = b \right\} \]

\[ S_{2^h} = \sum_{l=1}^{Q_{2^h}-1} P_l(l) \]  

\[ P_l(l) = \max_{b=0,1,\ldots,2^h-1} \left\{ \text{card} \{ k : k \in k_0, \ mod \ (k, 2^h) \} = b \right\} , \]  

where \( P_h(1) \leq P_h(2) \leq \cdots \leq P_h(2^{r-1}) \).

Signal independent uniqueness corresponds to the worst case signal form, when \( S_{2^h} = 0 \).

The answer is obtained almost immediately, since the computational complexity of the Theorem is of order \( O(N) \). The proof is given in [24].
For the considered set of missing samples \( n \in n_Q = \{0, 1, 2, 5, 7, 8, 11, 15, 18, 27, 28, 30\} \), from the previous example, we easily get

\[
Q_{a0} = 8,
Q_{a1} = \max\{6, 6\} = 6,
Q_{a2} = \max\{3, 2, 3, 4\} = 4,
Q_{a3} = \max\{2, 1, 2, 1, 1, 1, 1, 2\} = 4,
Q_{a4} = 2.
\]  

(23)

For signal independent uniqueness,

\[
s < 32 - \max\{1 \times (8 - 1), 2 \times (2 - 1), 4 \times (4 - 1), 8 \\
\times (2 - 1), 16 \times (2 - 1)\} - s = 16 - s.
\]

(24)

It means that \( s < 8 \), with considered \( n_Q \), guarantees uniqueness of the reconstructed signal for any signal parameters. When a sparse signal is reconstructed, for some sets of \( K \), uniqueness may be guaranteed for higher values of sparsity \( s \). Additional details and examples may be found in [24].
4. Statistical Analysis and Comparison

Consider a general real-valued form of a signal sparse in the DFT domain

\[ x(n) = \sum_{i=1}^{K} A_i \cos \left( \frac{2\pi n k_i}{N} + \phi_i \right). \]  

(25)

The total number of signal samples is \( N \). The signal sparsity \( s = 2K \) is varied from \( s = 2 \) to \( s = N/2 \). The signal amplitudes \( A_i \), frequencies \( k_i \), and phases \( \phi_i \) are random variables. Signal amplitudes \( A_i \) are Gaussian random variables with unity variance; the frequency indices \( k_i \) are random numbers within \( 1 \leq k_i \leq N - 1 \). The phases are uniform random variables \( 0 \leq \phi_i \leq 2\pi \), in each signal realization. Signal is reconstructed in 100 different realizations for each sparsity \( s \) and random positions of missing samples. The statistical calculations are performed for \( N = 128 \). The reconstructed signals are denoted by \( x_R(n) \). The reconstruction results are
Figure 3: Signal-to-reconstruction-error (SRR) obtained by using Algorithm 1, averaged over 100 realizations for various values of sparsity $s$ and number of available samples $M$. (a) The total number of samples is $N = 128$. (b) With a Gaussian noise in the input signal, SNR = 20 [dB] and $N = 64$.

The efficiency of the presented algorithm is compared with the standard routines used for the $\ell_1$-norm based minimization problems. The performance of the proposed algorithm are compared with the algorithm that recasts the recovery problem (3) into regression framework (LASSO-ISTA) and a linear programming framework with the primal-dual interior point method (L1-magic code in MATLAB). A simple review of the LASSO-ISTA algorithm is presented in Appendix. All there algorithms are run using 100 sparse signals with random parameters. The algorithm parameters are set in such a way that the average computational time for all cases is about 25 ms. The results are presented in Table 1. Columns’ notation in the table is $s$ for sparsity and $Q = N - M$ for the number of missing samples and MAE stands for the mean absolute error. Calculation time using MATLAB is presented in all cases.

We can conclude that in most of the considered cases the proposed algorithm outperforms other two algorithms in both the calculation time and accuracy.

Example 3. The algorithm can be applied to other signal transforms as well [23]. An example with missing (salt-and-pepper noise) pixels in an image with the 2D-DCT as its sparsity domain will be presented. Consider a standard MATLAB test image canoe.tif (Figure 5) and assume that 50% of the samples are corrupted (noisy image). These pixels with extreme values are removed and considered as unavailable. Since the coefficients in the 2D-DCT converge fast we will assume that the sparsity condition is satisfied (without imposing it in an explicit way by making some 2D-DCT coefficients equal zero). The image is reconstructed by minimizing the 2D-DCT sparsity measure and varying missing pixel values according to the presented algorithm.
Figure 4: (a), (b) The percentage of the full recovery events as a function of the number of available samples $M$ and the sparsity $s$ in the case of $N = 64$. (c) The average number of iterations as a function of the number of missing samples and sparsity. (d) The average computation time.

Table 1: MAE and computational time for the L1-magic (LP-DP), LASSO-ISTA, and the proposed algorithm.

<table>
<thead>
<tr>
<th>$s$</th>
<th>$N-M$</th>
<th>L1-magic</th>
<th>LASSO</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MAE</td>
<td>Time</td>
<td>MAE</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>$7.94 \times 10^{-11}$</td>
<td>29.7</td>
<td>$1.24 \times 10^{-4}$</td>
</tr>
<tr>
<td>16</td>
<td>8</td>
<td>$6.70 \times 10^{-10}$</td>
<td>28.4</td>
<td>$1.72 \times 10^{-4}$</td>
</tr>
<tr>
<td>32</td>
<td>8</td>
<td>$1.28 \times 10^{-9}$</td>
<td>29.8</td>
<td>$2.25 \times 10^{-4}$</td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>$1.64 \times 10^{-10}$</td>
<td>31.2</td>
<td>$1.30 \times 10^{-4}$</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>$1.22 \times 10^{-5}$</td>
<td>25.4</td>
<td>$1.91 \times 10^{-4}$</td>
</tr>
<tr>
<td>32</td>
<td>16</td>
<td>$5.44 \times 10^{-5}$</td>
<td>22.7</td>
<td>$7.16 \times 10^{-4}$</td>
</tr>
<tr>
<td>8</td>
<td>24</td>
<td>$2.83 \times 10^{-10}$</td>
<td>23.2</td>
<td>$1.37 \times 10^{-4}$</td>
</tr>
<tr>
<td>16</td>
<td>24</td>
<td>$4.62 \times 10^{-6}$</td>
<td>21.6</td>
<td>$2.51 \times 10^{-4}$</td>
</tr>
<tr>
<td>32</td>
<td>24</td>
<td>$2.58 \times 10^{-4}$</td>
<td>24.2</td>
<td>$1.17 \times 10^{-3}$</td>
</tr>
<tr>
<td>8</td>
<td>32</td>
<td>$6.66 \times 10^{-10}$</td>
<td>20.4</td>
<td>$1.51 \times 10^{-4}$</td>
</tr>
<tr>
<td>16</td>
<td>32</td>
<td>$1.94 \times 10^{-4}$</td>
<td>22.4</td>
<td>$4.09 \times 10^{-4}$</td>
</tr>
<tr>
<td>32</td>
<td>32</td>
<td>$1.08 \times 10^{-4}$</td>
<td>25.0</td>
<td>$8.08 \times 10^{-4}$</td>
</tr>
</tbody>
</table>
The original, corrupted, and reconstructed images after 4th and 50th iteration are shown in Figure 5.

5. Conclusion

Performance of the recently proposed gradient-based signal reconstruction algorithm in the domain of the signal samples/measurements is studied. Application of this algorithm is very efficient in signal processing problems with reduced set of data, since it is easy to define a set of missing signal samples/measurements. The values of signal samples are considered as variables, while the available samples are fixed and not changed in the algorithm. An efficient criterion, based on the gradient directions, is proposed for the algorithm step-size adaptation. Statistical analysis confirms calculation efficiency. The algorithm can easily be adapted to the cases on nonuniform signal samples, when only a reduced set of linear combinations of the signal samples are available.

Appendix

Gradient-Based Reconstruction in the Sparsity Domain

One of the forms for formulation of minimization problem (3), in order to find its solution, is based on the Lagrangian

$$ F(X) = \|y - AX\|_2^2 + \lambda \|X\|_1. $$

The minimization function is denoted by $F(X)$. The signal reconstruction in its sparsity domain is formulated as [9]

$$ X = \arg \min_X \{\|y - AX\|_2^2 + \lambda \|X\|_1\}. $$

This problem can be solved using least absolute selection and shrinking operator (LASSO) minimization. Parameter $\lambda$ defines a balance between the squared absolute error and sparsity.

The $l_1$-norm based reconstruction algorithms do not have a closed form solution. In order to define an iterative procedure a nonnegative term

$$ G(X) = (X - X_0)^T (\alpha I - A^T A)(X - X_0) $$

is added. It has a zero value at the position of the problem solution, denoted by $X_0$. This term is added to the function $F(X)$. New minimization function is

$$ H(X) = F(X) + (X - X_0)^T (\alpha I - A^T A)(X - X_0). $$

The solution follows from

$$ \nabla H(X) = \frac{\partial H(X)}{\partial X^T} = 0 $$

as

$$ X + \frac{\lambda}{2\alpha} \text{sign} (X) = \frac{1}{\alpha} A^T (y - AX_0) + X_0. $$

The iterative procedure can be defined by

$$ X_{n+1} + \frac{\lambda}{2\alpha} \text{sign} (X_{n+1}) = \frac{1}{\alpha} A^T (y - AX_n) + X_0. $$
The equation of the form
\[ x + \lambda \text{sign}(x) = y \] (A.8)
can be solved by using the soft-thresholding rule
\[ x = \text{soft}(y, \lambda) = \text{sign}(y) \max\{0, |y| - \lambda\}. \] (A.9)
The same rule can be used for each coordinate of the sparsity vector in (A.7). The iterative form for the signal reconstruction in the sparsity domain is
\[ X_{s+1} = \text{soft}\left(\frac{1}{\alpha} A^T (y - AX_s) + X_s, \frac{\lambda}{2\alpha}\right). \] (A.10)
This is the iterative soft-thresholding algorithm (ISTA) for problem (A.2) minimization [25]. The initial estimate is \( X_0 = A^T y \), corresponding to the transform of initial signal (4).

**Competing Interests**

The authors declare that they have no competing interests.

**References**


