Nonsparse influence on reconstruction with sparsity constraint

Isidora Stanković
GIPSA Lab, Grenoble INP, University of Grenoble Alpes, 38000 Grenoble, France
Faculty of Electrical Engineering, University of Montenegro, 81000 Podgorica, Montenegro
E-mail: isidora.stankovic@gipsa-lab.grenoble-inp.fr

I. ABSTRACT

A. Brief introduction

The theoretical results of this abstract, together with the proof of the theorem, are published in [1]. The application to audio signals was presented in [2]. The cases of random sampling, and with noisy samples, were considered in [3], [4].

A sparse signal is a signal with only few nonzero components in the transform domain. That signal can be reconstructed with less samples than using the traditional sampling theorem. The theory of sparse signal processing (SSP) and compressive sensing (CS) deals with the analysis and reconstruction of such signals.

In general, time-varying signals (such as audio signals) are not strictly sparse in the transformation domain. Because of their non-stationarity, the signals are either approximately sparse or not sparse. We say that a signal is $K$-sparse in a transformation domain if it has $K$ nonzero coefficients only and others are zero-valued. A signal is approximately sparse if the $K$ coefficients are significantly larger than other coefficients in the transform domain. A signal is said to be nonsparse if the $K$ coefficients of the signal are of the same order as other coefficients, meaning that the components of the signal are not distinguishable from other components in the signal. To use the theory of compressive sensing for any of these signals, the sparsity assumption has to be made.

In this research, we examine the influence of the non-reconstructed coefficients obtained by assuming that signals are reconstructed as $K$-sparse in the short-time Fourier transform (STFT) domain. Unlike other CS literature [5]–[7], an exact error of the reconstruction in the time-frequency domain is derived. The error depends on the number of the actual sparsity level and the number of available samples. It is validated by comparing it with the statistical reconstruction error.

B. Theorem

The reconstruction error of sparse signals is an important topic in compressive sensing. Its general bounds can be found in [5]–[7]. An exact formula for the expected squared reconstruction error, with the STFT as a sparsity domain, is presented by the next theorem.

**Theorem:** Consider a signal $x(n)$ with time-varying components. Its STFT values are denoted by $S_N(n) = [S_N(n, 0), S_N(n, 1), \ldots, S_N(n, N - 1)]^T$. The total number of signal samples within a window is $N$. Assume that the available signal samples are at $N_A$ random positions, defined by $n + m \in N_A$, and $N_M = N - N_A$ is the number of unavailable/missing samples. The signal is reconstructed under the assumption as it were $K$-sparse in the STFT domain (with the assumption that the reconstruction conditions are met for this sparsity). The reconstructed signal with $K$ nonzero STFT coefficients at $k \in \mathbb{K}$ is denoted by $S_{NR}(n)$. The error in the $K$ reconstructed STFT coefficients is:

$$
\|S_{NK}(n) - S_{NR}(n)\|^2 = K \frac{N_M}{N_A} \|S_N(n) - S_{NK}(n)\|^2. 
$$

(1)

The $K$-sparse version of $S_N(n)$ is denoted by $S_{NK}(n)$. The elements of vector $S_{NK}(n)$ are $S_{NK}(n, k) = S_N(n, k)$ for $k \in \mathbb{K}$ and $S_{NK}(n, k) = 0$ for $k \notin \mathbb{K}$. The reconstructed STFT $S_{NR}(n)$ is formed in the same way, with coefficients for $k \in \mathbb{K}$ being set to zero. Notation $\|S_N(n)\|^2$ is used for the expected value of the squared norm-two, i.e. $\|S_N(n)\|^2 = E[\sum_{k} |S_N(n, k)|^2]$.

The theorem is for the case of uniformly sampled measurements. For the case of a randomly sampled noisy signal, with noise variance being $\sigma_n^2$, the error is calculated as [3], [4]

$$
\|S_{NK}(n) - S_{NR}(n)\|^2 = \frac{K}{N_A} \|S_N(t_n) - S_{NK}(t_n)\|^2 + K \frac{N^2}{N_A} \sigma_n^2. 
$$

(2)

C. Results

**Example 1 - synthetic:** A combination of two linear frequency modulated signal components is considered. The reconstruction is done in an iterative way, using a variant of [8]. The STFT is calculated using a Hamming window of length $N = 256$ with a step in time of 32. The signal is not sparse in the DFT domain since its components sweep almost the whole frequency range. Various numbers of randomly positioned available samples $N_A$ have been considered, with sparsity levels $K = 4, 8, 16, 32, 64$.

The total reconstruction errors, statistical $E_{tot}$ and derived theoretical $E_{tot}$ obtained from Eq. (1), are calculated as

$$
E_{tot} = 10 \log \left( \|S_N(n) - S_{NR}(n)\|^2 \right) 
$$

(3)

$$
E_{tot} = 10 \log \left( \left( K \frac{N_M}{N_A} + 1 \right) \|S_N(n) - S_{NK}(n)\|^2 \right). 
$$

(4)

The total reconstruction error (averaged over 100 realizations) as a function of the number of available samples is presented in Fig. 1.

**Example 2 - recorded audio signal:** Now we will assume a recorded version of the words "You and I". It is assumed that there is only a half of the samples available. The assumed sparsity is $K = 75$. The original signal, the signal with available samples and the reconstructed one are shown in Fig. 2. The total error in dB caused by the reconstruction for various sparsity levels $K$ is shown in Fig. 3, with a very high agreement between the theory and estimation.

**Example 3 - randomly sampled with noise:** Let us consider a randomly sampled with three linear frequency modulated signal components, affected with zero-mean Gaussian noise with a standard deviation of $\sigma_n = 0.1$. We assume a Hann window of length $N = 256$ with a step in time of 32 is used for the reconstruction.
Using Eq. (3) for statistical error and

$$E_{\text{tot}}^\text{stat} = 10 \log \left( \frac{K}{N_A} + 1 \right) \| S_N(t_n) - S_{NK}(t_n) \|_2^2 + KN_A^2 \sigma^2 \epsilon \right),$$

(5)

for theoretical error, obtained from Eq. (2), Table I examines the total error of the reconstruction with various $K$ and $N_A$, averaged over 100 realizations.

**Keywords** – audio signals, error calculation, nonsparsity, sparse signal reconstruction, time-frequency analysis

<table>
<thead>
<tr>
<th>$K$</th>
<th>$N_A = N/2$</th>
<th>$2N/3$</th>
<th>$3N/4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1.006</td>
<td>1.147</td>
<td>0.932</td>
</tr>
<tr>
<td>16</td>
<td>0.103</td>
<td>0.298</td>
<td>0.034</td>
</tr>
<tr>
<td>32</td>
<td>−0.498</td>
<td>−0.210</td>
<td>−0.562</td>
</tr>
<tr>
<td>48</td>
<td>−0.480</td>
<td>−0.148</td>
<td>−0.572</td>
</tr>
</tbody>
</table>

Figure 1. Total reconstruction error as a function of the number of available samples $N_A$ for various assumed sparsity $K$. Theoretical results are presented by lines and the statistical with dots. Dots for $N_A > 4K$, when the reconstruction is possible with a high probability, are filled.

Figure 2. STFT reconstruction of the recorded audio signal “You and I”: Original STFT (top); STFT of the signal with available samples (middle); Reconstructed STFT (bottom).

Figure 3. Total error energy after the reconstruction with various sparsity levels of the recorded audio signal “You and I”. Blue solid line represents theory, red stars represents estimation (statistical values).

**REFERENCES**


