# ROBUST RECONSTRUCTION FOR CS-BASED FETAL BEATS DETECTION

*Giulia Da Poian, Riccardo Bernardini and Roberto Rinaldo*

DPIA, University of Udine, Udine, Italy

**ABSTRACT**

Due to its possible low-power implementation, Compressed Sensing (CS) is an attractive tool for physiological signal acquisition in emerging scenarios like Wireless Body Sensor Networks (WBSN) and telemonitoring applications. In this work we consider the continuous monitoring and analysis of the fetal ECG signal (fECG). We propose a modification of the low-complexity CS reconstruction SL0 algorithm, improving its robustness in the presence of noisy original signals and possibly ill-conditioned sensing/reconstruction procedures. We show that, while maintaining the same computational cost of the original algorithm, the proposed modification significantly improves the reconstruction quality, both for synthetic and real-world ECG signals. We also show that the proposed algorithm allows robust heart beat classification when sparse matrices, implementable with very low computational complexity, are used for compressed sensing of the ECG signal.

***Index Terms*—** Non-invasive Fetal ECG, Compressive sensing, Sparse representations.

## 1. INTRODUCTION

The Compressive Sensing paradigm asserts that one can successfully recover certain signals, sampled far below the Nyquist frequency, given that they are sparse in some dictionary [1] [2]. CS takes advantage of sparsity to reconstruct the compressed signal at the decoder/receiver side. In particular, an $N-$dimensional vector of signal samples is represented as a *sparse* linear combination of $K$ dictionary vector *atoms*, with $K \ll N$. By taking a small number $M > K$ of linear projections of the signal along random vectors, arranged in the so called *sensing matrix*, it is possible to reconstruct the signal exactly under certain circumstances [1], by minimizing the $l0$ norm of the representation, which corresponds to the number of its non zero elements. Due to the promising low power implementation of CS based compression schemes, the technique is particularly suitable for physiological signal telemonitoring via a wireless body-area network (WBAN), using low complexity battery operated devices. In particular, non-invasive continuous fetal electrocardiogram (fECG), recorded in a non invasive way from the maternal abdomen, can provide early detection of fetal arrhythmias, making it possible to perform a timely medical treatment. Moreover, the sparse signal structure exploited by CS for compression and transmission may also be used for heart beat detection and classification. Raw abdominal fECG signals are usually contaminated by many sources of noise and result to be non-sparse, even when represented with Wavelets or DCT bases, which are commonly used for the sparsification of adult ECG. However, it has been experimentally demonstrated that specifically designed dictionaries can provide a sparser representation [3], [4].

Starting from the CS measurements, common reconstruction algorithms [5], [6] relax the original NP-hard problem using approximations of the $l0$ norm of the sparse vector. The Smoothed-$L0$ algorithm presented in [6] (SL0) is particularly interesting for its low complexity, which allows for real-time signal reconstruction [7]. However, as we will see, these approaches may fail in case of noisy signals or ill-conditioned sensing/dictionary-based reconstruction procedures.

In this work we describe a variant of the SL0 technique for this more challenging setting. We show that, while maintaining the same computational cost of the original algorithm, the proposed modification significantly improves the reconstruction quality, both for synthetic and real-world ECG signals. We also show that the proposed algorithm allows robust heart beat classification when sparse matrices, implementable with very low computational complexity, are used for compressed sensing of the ECG signal.

## 2. REGULARIZATION OF SMOOTHED $L0$ ALGORITHM

Compressive Sensing aims to reconstruct a signal $\mathbf{x} \in \mathbb{R}^N$, sparse in some domain, by solving the following optimization problem

$$\min_{s} ||\mathbf{s}||_0 \text{ s.t. } \| \mathbf{y} - \mathbf{A}\mathbf{s} \|_2 \leq \epsilon, \quad \mathbf{A} = \mathbf{\Phi}\mathbf{D}, \qquad (1)$$

where $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$ is the sensing matrix, $\mathbf{x} = \mathbf{D}\mathbf{s}$ is the signal to be reconstructed and $\mathbf{y} = \mathbf{\Phi}\mathbf{x}$ is the vector of measurements. In (1), $\mathbf{D} \in \mathbf{R}^{N \times D}$ is the sparsifying dictionary and $\mathbf{s} \in \mathbb{R}^D$ is the sparse representation of $\mathbf{x}$, ideally with a small number $K$ of non-zero elements. The power consumption of CS based compression is strongly affected by the sensing matrix. While strong theoretical results suggest the use of a matrix with entries drawn as independent Gaussian random variables [2], sparse binary matrices, where $\mathbf{\Phi}$ has only $d$ non-zero randomly selected entries in each column, have been proposed to reduce the computational cost [8]. In this case, calculating $\mathbf{\Phi x}$ takes only $O(dN)$ operations, with a significant saving when $d \ll N$.

The SL0 algorithm proposed in [6] solves the problem in Eq. (1) by approximating the $l0$-norm with a continuous function, and optimizing the resulting cost function to provide a smooth measure of sparsity. Indeed, the $l0$-norm can be approximated using Gaussian functions, for small $\sigma$ values [6], as in

$$||\mathbf{s}||_{S,0} \triangleq D - \sum_{i=1}^{D} \exp(-s_i^2/2\sigma^2). \tag{2}$$

Thus, the minimization of the $l0$-norm is approximately equivalent to maximize $F_\sigma(s) = \sum_i \exp(-s_i^2/2\sigma^2)$. This enables to replace the $l0$-norm minimization with a convex problem, and maximize $F_\sigma(s)$ using a steepest ascent algorithm. The parameter $\sigma$ controls the trade-off between the smoothness of the objective function and the accuracy of the approximation of the $l0$-norm.

The algorithm proposed in [6] consists of two nested iterations, and the external loop is responsible to gradually decrease the $\sigma$ value. Note that, when $\sigma$ is sufficiently large, $\exp(-s_i^2/2\sigma^2) \approx 1 - s_i^2/2\sigma^2$, and the maximization of $F_\sigma(s)$ s.t. $\mathbf{y} = \mathbf{As}$ resembles the minimum $l2$-norm solution of $F_\sigma(s)$ s.t. $\mathbf{y} = \mathbf{As}$ [6]. Therefore, the starting solution of the optimization process is usually calculated using the pseudo-inverse $\mathbf{A}^\dagger$ of $\mathbf{A}$ and set to $\mathbf{s_0} = \mathbf{A}^T(\mathbf{AA}^T)^{-1}\mathbf{y}$.

The internal loop tries to maximize $F_\sigma(s)$ on the feasible set $\{\mathbf{s}|\mathbf{y} = \mathbf{As}\}$, using a steepest ascent algorithm, and updating $\mathbf{s} \leftarrow \mathbf{s} - \mu\delta_k$ where

$$\delta_k = \mathbf{s} \cdot \left[ e^{-\frac{s_1^2}{2\sigma_k^2}}, \ldots, e^{-\frac{s_D^2}{2\sigma_k^2}} \right]^T. \tag{3}$$

The next step consists in projecting $\mathbf{s}$ into the convex set to avoid trapping the algorithm in local maxima

$$\mathbf{s} = \mathbf{s} - \mathbf{A}^T(\mathbf{AA}^T)^{-1}(\mathbf{As} - \mathbf{y}). \tag{4}$$

The SL0 is typically 2 to 3 times faster than the Basis Pursuit denoising algorithm (SPGL1 implementation, [5]), while resulting in many cases in the same or better accuracy [6]. Note that the algorithm requires that matrix $\mathbf{A}$ has full rank $M$. When the number of measurements $M$ increases (i.e., the compression ratio decreases), this requirement may become critical, making the reconstruction problem ill-conditioned and sensitive to noise. To avoid these problems, in the next section we propose a regularized version of the reconstruction algorithm.

### 2.1. Regularization

In real scenarios where the sparse signal or the measurements are affected by noise, if the compound matrix $\mathbf{A}$ is ill-conditioned, then application of $\mathbf{A}^\dagger$ amplifies the error and results in poor reconstruction, even using the Robust SL0 proposed in [9]. Introducing a regularization term in the optimization problem enables a stable recovery of $\mathbf{x} = \mathbf{Ds}$.

As in the SL0 algorithm, we approximate the $l0$-norm by using (2), and the algorithm again consists in two nested iterations. The internal loop seeks the maximum of $F_\sigma$ in the feasible set $\{\mathbf{s}| \parallel \mathbf{y} - \mathbf{As} \parallel_2 \leq \epsilon\}$. At each step we compute $\tilde{\mathbf{s}} = \mathbf{s} - \mu\delta_k$ and project $\tilde{\mathbf{s}}$ by solving

$$\min_{\hat{s}} \parallel \hat{\mathbf{s}} - \tilde{\mathbf{s}} \parallel_2 \quad \text{s.t.} \quad \parallel \mathbf{A}\hat{\mathbf{s}} - \mathbf{y} \parallel_2 \leq \epsilon. \tag{5}$$

Using the Lagrangian function of Eq. (5), the problem can be rewritten as

$$\min_{\hat{s}} \parallel \mathbf{A}\hat{\mathbf{s}} - \mathbf{y} \parallel_2^2 + \lambda \parallel \hat{\mathbf{s}} - \tilde{\mathbf{s}} \parallel_2^2, \tag{6}$$

where $\lambda$ is the regularization parameter. The solution is

$$\hat{\mathbf{s}} = \tilde{\mathbf{s}} - \mathbf{A}^T(\mathbf{AA}^T + \lambda\mathbf{I}_M)^{-1}(\mathbf{A}\tilde{\mathbf{s}} - \mathbf{y}). \tag{7}$$

As for the SL0 algorithm, for large $\sigma$ values, the solution is equal to the $l2$ norm solution subject to $\parallel \mathbf{y} - \mathbf{As} \parallel_2 \leq \epsilon$. Solving the problem

$$\min_{s} \parallel \mathbf{As} - \mathbf{y} \parallel_2^2 + \lambda \parallel \mathbf{s} \parallel_2^2, \tag{8}$$

we set the initial solution of the algorithm to $\mathbf{s_0} = \mathbf{A^T}(\mathbf{AA^T} + \lambda\mathbf{I}_M)^{-1}\mathbf{y}$.

The proposed $\lambda$SL0 algorithm is summarized in Algorithm 1. The value of the regularization parameter $\lambda$ represents a compromise between the two terms of the cost function. When the noise norm $\epsilon$ is small, $\lambda \to 0$, and the algorithm reduces to the original SL0 for the noiseless case. We carried out some experiments (results are omitted due to lack of space) and we observed that the value of $\lambda$ is not critical and should be $\sim$10$-$100 times the expected noise $\epsilon$.

**Algorithm 1** $\lambda$SL0

**Input:** $\mu$ step size, $\mathbf{y}$, $\mathbf{A}$, $\sigma_{dec}$, $\sigma_{min}$, $\lambda$, $K_{iter}$
**Initialization:** $\mathbf{s}_0 \leftarrow \mathbf{A^T}((\mathbf{AA^T}) + \lambda\mathbf{I}_M)^{-1}\mathbf{y}$,
  $\sigma_1 = 2|\max(\mathbf{s}_0)|$
  **while** $\sigma_k < \sigma_{min}$ **do**
    **for** k=1:K$_{iter}$ **do**
      $\delta_k \leftarrow \mathbf{s} \cdot [e^{-\frac{s_1^2}{2\sigma_k^2}}, \ldots, e^{-\frac{s_D^2}{2\sigma_k^2}}]^T$
      $\mathbf{s} \leftarrow \mathbf{s} - \mu\delta_k$
      Project $\mathbf{s}$ onto the feasible set: $\{\mathbf{s}| \parallel \mathbf{As} - \mathbf{y} \parallel_2 \leq \epsilon\}$
      $\mathbf{s} \leftarrow \mathbf{s} - \mathbf{A^T}((\mathbf{AA^T}) + \lambda\mathbf{I}_M)^{-1}(\mathbf{As} - \mathbf{y})$
    **end for**
    $\sigma_k \leftarrow \sigma_k\sigma_{dec}$
    $\tilde{\mathbf{s}}_\mathbf{k} \leftarrow \mathbf{s}$
  **end while**
**Output:** $\mathbf{s_{OUT}} \leftarrow \tilde{\mathbf{s}}_\mathbf{k}$

## 3. PERFORMANCE OF $\lambda$SL0

In this section, the effect of noise on the reconstruction performance is experimentally analyzed. We compare the performance of the proposed algorithm with the original SL0 and the BPDN-SPGL1 algorithms. The signals used in these experiments are simulated fECG signals [10] with length $N = 256$. As sparsifying dictionaries we use a dictionary of Gaussian like functions [7], and the Wavelet basis with Daubechies' length-4 filters. The sensing matrix elements are drawn as independent Gaussian random variables [2]. We repeat the experiment 100 times with different source signals at different noise levels, and using each time a different random sensing matrix. The reported SNR value is the average of these simulations. In Fig. 1, we report the reconstruction SNR as a function of the input SNR (SNRin) when gaussian noise is added to the simulated fECG traces, for a compression ratio CR=50% ($M = 128$). Compared to the original SL0 algorithm, $\lambda$SL0 allows to achieve better reconstruction quality, especially when the Gaussian Dictionary [7] is used. An improvement can be also appreciated when the Wavelet basis is used, especially at lower SNRin values. Note that the use of the Gaussian Dictionary gives much better performance than Wavelets also when BPDN is used.

In addition to the previous experiments, we assess the reconstruction performance as the compression ratio changes. In Fig. 2, it is possible to see that the average SNR achieved by the $\lambda$SL0 algorithm combined with the Gaussian Dictionary outperforms the SL0 method, especially at low compression ratios ($M$ large) and is comparable with respect to the traditional BPDN algorithm, which has a higher complexity. At higher compression ratios (CR$> 50\%$) the Wavelet basis achieves a lower performance independently of the reconstruction algorithm.

As a measure of the computational cost of the algorithms we use the CPU time, setting the same parameters ($K_{iter} = 3$ an $\sigma_{dec} = 0.5$ ) for SL0 and $\lambda$SL0. Experiments show an average reconstruction time for the $\lambda$SL0 algorithm ranging from 0.07 s, when CR=30%, to 0.01 s, when CR=80%. Thus, it maintains approximately the same computational cost of the original SL0 algorithm (ranging from 0.03 s to 0.01 s), while being much faster than the BPDN algorithm (1.6 s to 0.6 s). Programs are written in Matlab, running on an Intel Core i7 processor, equipped with 16 GB memory.

## 4. APPLICATION TO JOINT COMPRESSION AND BEATS DETECTION IN FECG

In this section, we analyze the performance of $\lambda$SL0 for fetal beat detection in CS-compressed real-world fECG signals. We also analyze the influence of different sensing matrices, in particular sparse matrices which allow very low complexity of the CS sensor.

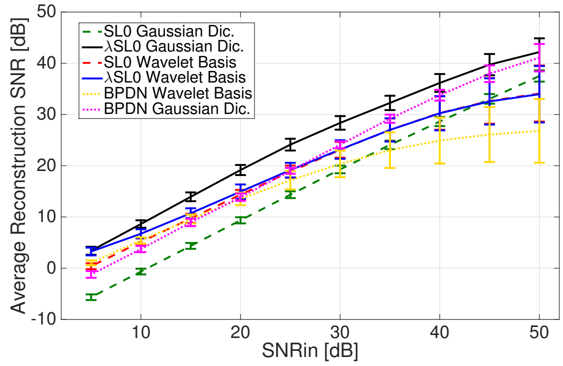



**Fig. 1**. Reconstruction SNR versus input SNR obtained from 100 trials for simulated fECG signals, at CR=50%, using the SL0, $\lambda$SL0 and BPBN (SPGL1) algorithms using the Wavelet and the Gaussian Dictionary.

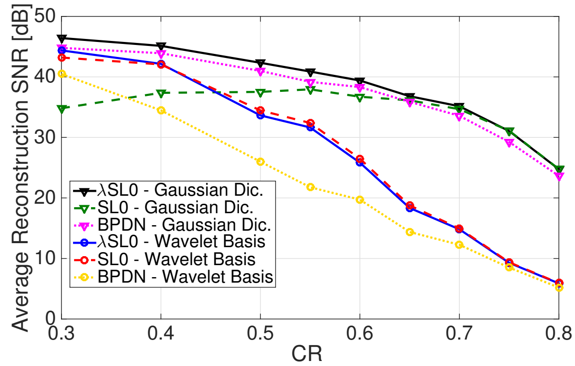



**Fig. 2**. Reconstruction SNR versus CR obtained from 100 trials for simulated fECG signals.

### 4.1. fECG Reconstruction and Fetal Beats Detection

In [7] a framework for the compression of multichannel abdominal fECG and joint detection of fetal beats has been proposed. The compression of the signal is based on Compressive Sensing and uses a binary sparse sensing matrix, containing only $d = 2$ ones in random positions in each column, in order to reduce the sensor complexity [8]. Before reconstruction using SL0, Independent Component Analysis (ICA) is applied on the compressed measurements, and then the reconstruction process recovers the uncompressed independent components (ICs). The sparse decomposition used to reconstruct the ICs is also used to further separate the maternal and fetal signals and to detect the time location of the beats. In particular, we use a dictionary of Gaussian like functions [4], composed by two sub-dictionaries for the approximation of the maternal ECG component and the fetal one. The separation is based on the atoms, belonging to the fetal or mother's sub-dictionaries, activated during the reconstruction process. In accordance with the analysis of the previous section,we found out experimentally that, for compression ratios greater than 50%, the detection performance is preserved, while at lower compression ratios, besides the increased information available, fetal beat detection may fail.

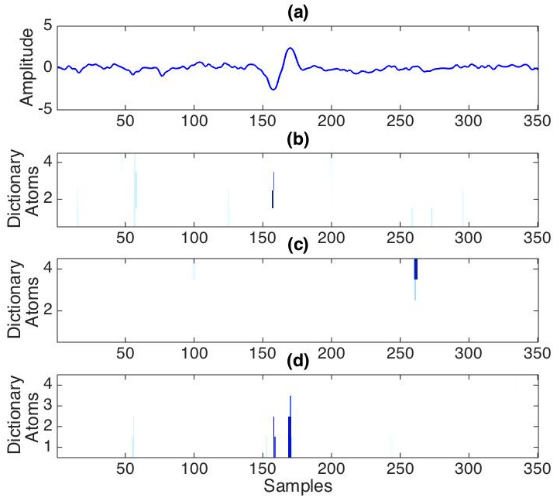
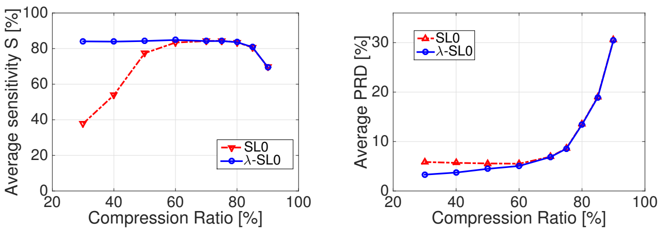


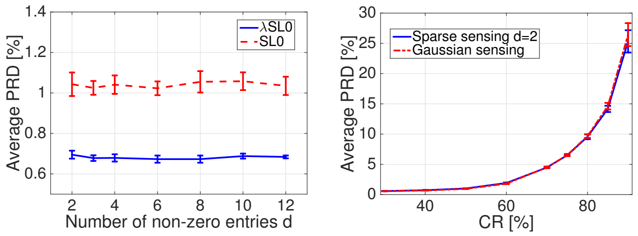

**Fig. 3**. Sparse decomposition of the independent component in (a) using the SL0 algorithm, for (b) CR=75% and (c) CR=40% and (d) using the $\lambda$SL0 algorithm for CR=40%. In the graphs, different intensities represent the weight of the activated atoms.




(a) (b)

**Fig. 4**. (a) Detection performance for SL0 and $\lambda$SL0 algorithm. The vertical coordinate gives the average Sensitivity for dataset A at different CR values. (b) Comparison of average PRD when using the two algorithms at different CRs.

As an example, we show in Fig. 3 (a) a portion of the IC of signal a32 of the 4-channel Physionet Challenge dataset [11], where the fetal beat is clearly visible. When the compression ratio is CR=75%, the reconstruction quality value PRD (percentage root-mean-square difference) is about 6% (average of the 4 channels) and the detection performance in terms of sensitivity (S) and positive predictivity (P+) are S=100% and P+=99.34%, respectively. At lower compression ratios detection fails, and for CR=40% we have PRD=0.47%, S=60% and P+=68%. Fig. 3 (b) and (c) show the positions of the fetal activated dictionary atoms for the two cases CR=75% and CR=40%, respectively. As we can see, when CR=40%, the algorithm fails to find the correct sparse representation, leading to a wrong beat detection. Instead, Fig. 3 (d) shows the




(a) (b)

**Fig. 5**. Average recovery quality for signal a25 of dataset A. The vertical coordinate gives the average PRD and the error bar gives the standard deviation. (a) Effects of the number of non-zero entries in each column of the sensing matrix at CR=40% for SL0 and $\lambda$SL0. (b) Comparison of average PRD using sparse sensing matrices with $d = 2$ and random Gaussian sensing matrices at different CR ($\lambda$SL0 algorithm).

activated atoms when the $\lambda$SL0 algorithm is used instead of SL0. For the whole signal and using the same sensing matrix at CR=40%, $\lambda$SL0 achieves S=100% and P+=99%. We repeat the experiment 20 times with different random sparse binary matrix ($d = 2$), for all the signals in dataset A (excluding badly annotated signal as in [12]). The reported sensitivity value is the average of these simulations. Fig. 4 (a) shows that the detection performance of the proposed algorithm is almost independent of the CR, while the SL0 algorithm fails at lower CRs.
In Fig. 4 (b) it can be seen that the proposed $\lambda$SL0 algorithm outperforms the original algorithm, in terms of average reconstruction quality PRD, at low compression ratios.

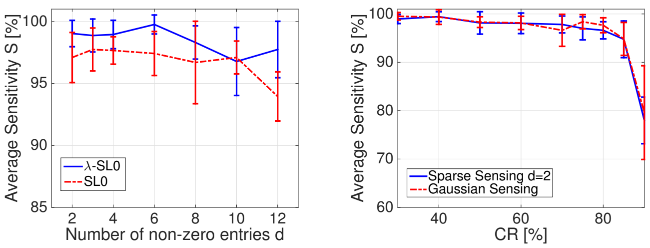
### 4.2. Influence of the sensing matrix

In this section, we analyze the performance of the reconstruction procedure, using SL0 and $\lambda$SL0 algorithms, when different sensing matrices, generated from i.i.d. Gaussian random variables or sparse with different $d$ values, are used. There are not theoretical guidelines for choosing the optimal number of non-zero elements $d$, therefore it has been determined experimentally. In the following experiments, the signal is divided into $N = 250$ sample long segments, which are compressed independently. For each $d$ value and compression ratio, the experiments were repeated using 20 randomly generated different sensing matrices, and the average performance is reported. Results for signal a25 of the Challenge dataset A are shown in Fig. 5 and Fig. 6. As we can see from Fig. 5(a), for a compression ratio CR=40%, increasing the number of ones in each column of the sensing matrix, does not improve the reconstruction quality, both for the SL0 and $\lambda$SL0 reconstruction algorithms. Note however that $\lambda$SL0 outperforms SL0. Fig. 5(b) shows a comparison of the reconstruction quality obtained with sparse matrices, $d = 2$, with respect to the quality obtained using Gaussian random matrices. Although the theo-

**Table 1**. Average performance of detection and reconstruction for SL0 and $\lambda$SL0 for dataset A.

| CR % | | SL0 S [%] | SL0 PRD [%] | $\lambda$SL0 S [%] | $\lambda$SL0 PRD [%] |
|---|---|---|---|---|---|
| 40 | Sparse 2 | 46 | 5.27 | 85 | 3.77 |
| | Gaussian | 45 | 5.58 | 84 | 3.72 |
| 50 | Sparse 2 | 77 | 5.93 | 85 | 4.49 |
| | Gaussian | 75 | 5.71 | 85 | 4.27 |
| 75 | Sparse 2 | 84 | 8.81 | 84 | 8.14 |
| | Gaussian | 84 | 8.80 | 84 | 8.02 |

retical reconstruction performance for i.i.d. Gaussian sensing matrices is well established, we can see experimentally that, for the class of signals we are considering, sparse matrices have similar performance. The use of a sparse sensing matrix with $d = 2$ allows to achieve almost identical reconstruction results, besides the very low complexity implementation. Finally, Table 1 summarizes the average reconstruction and detection performance for dataset A, at different compression ratios, when using a sparse sensing matrix with $d = 2$ and an i.i.d. Gaussian sensing matrix. Experiments are repeated for the SL0 and $\lambda$SL0 algorithms. Both detection and reconstruction are mostly independent from the sensing matrix, while it is apparent that the $\lambda$SL0 algorithm allows robust detection and reconstruction.




(a) (b)

**Fig. 6**. (a) Effects of the number of non-zero entries in each column of the sensing matrix on detection performance. The vertical coordinate gives the average Sensitivity S for signal a25 of dataset A for CR=40%, the error bar gives the standard deviation. (b) Comparison of average Sensitivity when using sparse sensing matrices with $d = 2$ and random Gaussian sensing matrices at different CR ($\lambda$SL0 algorithm).

## 5. CONCLUSIONS

In this paper we proposed a regularized version of the SL0 algorithm. Experimental results confirm that the proposed algorithm has good performance, while preserving the low computation complexity of the original one. The application of $\lambda$SL0 to the joint compression and detection framework of fetal ECG also demonstrates that the proposed modification can efficiently reconstruct the signals and correctly detect the beats in the presence of noise and for different compression ratios. Moreover, we have shown that the use of sparse sensing matrices with only 2 non-zero elements in each column, compares successfully with random Gaussian matrices, while permitting a very low complexity implementation.

## REFERENCES

[1] D. L. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] E. J. Candès *et al.*, "Compressive sampling," in *Proceedings of the international congress of mathematicians*, vol. 3. Madrid, Spain, 2006, pp. 1433–1452.

[3] D. Craven *et al.*, "Compressed sensing for bioelectric signals: A review," *Biomedical and Health Informatics, IEEE Journal of*, vol. 19, no. 2, pp. 529–540, 2015.

[4] G. Da Poian *et al.*, "Gaussian dictionary for compressive sensing of the ECG signal," in *Biometric Measurements and Systems for Security and Medical Applications (BIOMS) Proceedings, 2014 IEEE Workshop on*. IEEE, 2014, pp. 80–85.

[5] E. van den Berg and M. P. Friedlander, "Probing the pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2008.

[6] H. Mohimani *et al.*, "A fast approach for overcomplete sparse decomposition based on smoothed norm," *Signal Processing, IEEE Transactions on*, vol. 57, no. 1, pp. 289–301, 2009.

[7] G. Da Poian *et al.*, "Separation and analysis of fetal-ecg signals from compressed sensed abdominal ecg recordings," *Biomedical Engineering, IEEE Transactions on*, vol. PP, no. 99, pp. 1–11, 2015.

[8] R. Berinde and P. Indyk, "Sparse recovery using sparse random matrices," *MIT-CSAIL Technical Report*, 2008.

[9] A. Eftekhari *et al.*, "Robust-SL0 for stable sparse representation in noisy settings," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*. IEEE, 2009, pp. 3433–3436.

[10] J. Behar *et al.*, "An ECG simulator for generating maternal-foetal activity mixtures on abdominal ecg recordings," *Physiological measurement*, vol. 35, no. 8, p. 1537, 2014.

[11] "Physionet challenge 2013," http: www.physionet.org/challenge/2013/.

[12] J. Behar *et al.*, "Combining and benchmarking methods of foetal ECG extraction without maternal or scalp electrode data," *Physiological Measurement*, vol. 35, no. 8, p. 1569, 2014.