

# Structured Sparsity Models for Compressively Sensed Electrocardiogram Signals: A Comparative Study

Hossein Mamaghanian, Nadia Khaled, David Atienza and Pierre Vanderghenst  
 School of Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL)  
 Email: name.surname@epfl.ch

**Abstract**—We have recently quantified and validated the potential of the emerging compressed sensing (CS) paradigm for real-time energy-efficient electrocardiogram (ECG) compression on resource-constrained sensors. In the present work, we investigate applying sparsity models to exploit underlying structural information in recovery algorithms. More specifically, re-visiting well-known sparse recovery algorithms, we propose novel model-based adaptations for the robust recovery of compressible signals like ECG. Our results show significant performance gains for the recovery algorithms exploiting the underlying sparsity models.

## I. INTRODUCTION

CS is a new sensing and processing paradigm, which challenges the traditional analog-to-digital conversion based on the Shannon sampling theorem. For sparse signals such as the electrocardiogram (ECG), Nyquist-rate sampling produces a large amount of redundant digital samples, which are costly to wirelessly transmit in the context of our target mobile ECG monitoring systems, and require to be further compressed using non-linear digital techniques. CS is a methodology that has been recently proposed to address this problem.

Capitalizing on this sparsity, we have recently proposed [1], to apply the emerging compressed sensing (CS) approach [2] for a low-complexity, real-time and energy-efficient ECG signal compression on wireless body sensor network (WBSN) motes. We have also quantified the potential of the emerging compressed sensing (CS) signal acquisition/compression paradigm for low-complexity energy-efficient ECG compression on the state-of-the-art Shimmer<sup>TM</sup> WBSN mote. Interestingly, our results show that CS represents a competitive alternative to state-of-the-art digital wavelet transform (DWT)-based ECG compression solutions in the context of WBSN-based ECG monitoring systems. The results validates the suitability of compressed sensing for real-time energy-aware ECG compression on resource-constrained WBSN motes.

Recent works about practical utilization of CS for biomedical signal analysis confirm that in order to be competitive with other state-of-the-art compression algorithms, new CS methods that fully leverage the underlying structural information of the considered signals (beyond simple sparsity analysis) must be developed. Building on our previous work, we herein investigate state-of-the-art CS recovery algorithms and apply model-based recovery for ECG signals.

*Notation:* In all the following, normal letters designate scalar quantities, boldface lower-case letters indicate column

vectors, and boldface capitals represent matrices. Moreover,  $m_i$  and  $M_{i,j}$  are the  $i^{th}$  entry of vector  $\mathbf{m}$  and the  $(i,j)^{th}$  entry of matrix  $\mathbf{M}$ , respectively. Finally,  $(\cdot)^H$  and  $\|\cdot\|_p$  denote the conjugate transpose, and the  $l_p$ -norm of a vector, respectively.

## II. COMPRESSED SENSING AND SPARSE RECOVERY

Let  $\mathbf{x}$  be the real-valued  $N$ -dimensional ECG signal vector ( $\mathbf{x} \in \mathbb{R}^N$ ). As aforementioned, the original ECG signal  $\mathbf{x}$  has a sparse approximation, i.e., it can be represented by a linear superposition of  $S$  elements of an orthonormal wavelet basis,  $\mathbf{x} \approx \sum_{k=1}^S \alpha_k \psi_k$ , with  $S \ll N$ . Conventionally, one would collect ECG samples at the Nyquist rate forming  $\mathbf{x}$  and then compress it using non-linear digital compression techniques. CS offers a striking alternative by showing that you can collect roughly  $S$  samples using simple *analog* measurement waveforms, thus sensing/sampling and compressing at the same time. Accordingly, we collect  $M$  samples using simple measurement vectors  $\{\phi_i\}_{1 \leq i \leq M}$  as  $y_i = \phi_i^H \mathbf{x} = \langle \phi_i, \mathbf{x} \rangle$ ,  $i = 1, \dots, M$ . Consequently, the CS linearly compressed data vector  $\mathbf{y} \in \mathbb{R}^M$  is described by  $\mathbf{y} = \Phi \mathbf{x}$ , where  $\Phi$  denotes the  $M \times N$  measurement or sensing matrix with the vectors  $\phi_1^H, \dots, \phi_M^H$  as rows. It is important to notice that the sensing matrix  $\Phi$  does not depend on the signal: CS proposes a simple linear sampling strategy that is only marginally off the optimal but complex best adaptive strategy. To guarantee the robust and efficient recovery of the  $S$ -sparse signal  $\alpha_S$ , the sensing matrix  $\Phi$  must obey the key *restricted isometry property* (RIP) [3]:

$$(1 - \delta_S) \|\alpha\|_2 \leq \|\Phi \Psi \alpha\|_2 \leq (1 + \delta_S) \|\alpha\|_2, \quad (1)$$

for all  $S$ -sparse vectors  $\alpha$ .  $\delta_S$  is the isometry constant of matrix  $\Phi$ , which must be not too close to one. In words, the RIP ensures that all sub-matrices of  $\Phi \Psi$  of size  $M \times S$  are close to an isometry, and therefore preserve distance (and information). A universal good choice for the sensing matrix  $\Phi$  are random matrices, which are largely incoherent with any fixed basis  $\Psi$  [4]. By extension, random matrices with independent identically distributed (i.i.d.) entries formed by sampling (1) a Gaussian distribution  $\mathcal{N}(0, 1/N)$ ; (2) a symmetric Bernoulli distribution ( $P(\Phi_{i,j} = \pm 1/\sqrt{N}) = 1/2$ ) will satisfy the RIP with high probability with any orthonormal basis  $\Psi$  when  $M = \mathcal{O}(S \log(S/N))$ .

If RIP holds, then an approximate sparse signal reconstruction can be accomplished by solving the following convex optimization problem:

$$\min_{\tilde{\alpha} \in \mathbb{R}^N} \|\tilde{\alpha}\|_1 \quad \text{subject to} \quad \|\Phi\Psi\tilde{\alpha} - \mathbf{y}\|_2 \leq \sigma, \quad (2)$$

where  $\sigma$  bounds the amount of noise corrupting the data. In the context of this paper, we consider solving (2) using three different well-known sparse recovery algorithms: Iterative hard-thresholding (IHT), fast Lipschitz iterative hard-thresholding (FLIHT), and Nesterov's iterative hard-thresholding (NIHT). The last two algorithms are collectively referred to as algebraic pursuits algorithms (ALPS). These algorithms are briefly discussed in the sequel.

1) *Iterative Hard-Thresholding (IHT)*: is a popular method for sparse recovery in the compressed sensing literature [5]. IHT has the following recursion:  $\alpha_{i+1} = \alpha_i + \Phi^T(\mathbf{y} - \Phi\alpha_i)$ . Despite its extremely easy implementation, which makes it an alternative to convex optimization approaches for large-scale problems, IHT has a quite slow convergence rate [6]. In practice, in many cases for compressible signals, it is not uncommon to lose the convergence and quickly be unstable. This is often due to the selection of wrong coefficients in the hard-thresholding step.

2) *Algebraic Pursuits algorithms (ALPS)*: The two algebraic pursuit recovery algorithms used in this work are: (1) Fast Lipschitz Iterative Hard Thresholding (FLIHT) and (2) Nesterov's Iterative Hard Thresholding (NIHT). On the one hand, FLIHT is an optimization recovery algorithm which uses the first order scheme introduced by Nesterov (1983) in the recursion step [7]. Its convergence rate is proved to match the theoretical limit for the class of Lipschitz continuous and strongly convex functions [6]. On the other hand, NIHT is based on another gradient method proposed by Nesterov in 2003. Compared to FLIHT, NIHT incorporates a weighted history of past gradients. A detailed description of these algorithms can be found in [6].

### III. ENHANCED SPARSE RECOVERY BASED ON STRUCTURED SPARSITY MODELS

#### A. Structured Sparsity Models

While many natural and man-made signals and images can be described to the first-order as sparse or compressible, their sparse supports (set of nonzero coefficients) often have an underlying order. This order plays a central role in the transform compression literature, but it has barely been explored in the CS context [8]. New contention in recent works is that for CS to truly live up to its promise it must more fully leverage concepts from state-of-the-art compression algorithms. In virtually all such algorithms, the key ingredient is a signal model that goes beyond simple sparsity by providing a model for the basis coefficient structure. Exploiting underlying structural informations of coefficients vector in addition to the sparse representation prior in recovery problems potentially reduce the degrees of freedom of possible solutions. This permits only certain configurations of possible answers, which would

potentially lead to: (1) better quality of reconstructed signals by enabling us to better differentiate true signal information from recovery artifacts; (2) stable recovery using fewer measurements  $M$ ; and (3) faster reconstruction algorithms. The idea of structured or model-based CS has given rise to the design and development of sophisticated compression algorithms that operate on a given signal  $\mathbf{x}$  according to structured sparsity models.

As aforementioned, the support of their non-zero/large coefficients often has an underlying interdependency structure. A new framework for CS captures such structures using a *union-of-subspaces* model  $\mathcal{M}_k$  [9]. Let  $\mathbf{x}$  be an  $S$ -sparse signal with  $S \ll N$ . The set of indices corresponding to the nonzero entries are called, support of  $\mathbf{x}$  and denoted by  $\text{supp}(\mathbf{x})$ . The set of all  $S$ -sparse signals is the union of the  $\binom{N}{S}$ ,  $S$ -dimensional subspaces aligned with the coordinate axes in  $\mathbb{R}^N$ . This union of subspaces, denoted by  $\Sigma_S$ . Other than its  $S$ -sparsity, there are no further constraints on the support or values of its coefficients. A *union-of-subspaces* signal model endows the  $S$ -sparse signal  $\mathbf{x}$  with additional structure that allows certain  $S$ -dimensional subspaces in  $\Sigma_S$  and disallows others [9]. This will relax the RIP constraint on the CS measurement matrix  $\Phi$ . The equation 1 is required to hold only for signals  $\mathbf{x} \in \mathcal{M}_k$ . This new definition of the of the RIP constraint is called  $\mathcal{M}_k$ -RIP property [5].

Given the highly sparse nature of the ECG signals in wavelet domain, it is used as our sparse representation matrix  $\Psi$ .

#### B. Wavelet Tree Model for ECG Signal Recovery

Wavelet coefficients can be naturally organized into a tree structure, and for many kinds of natural and manmade signals the largest coefficients cluster along the branches of this tree. Consider a signal  $\mathbf{x}$  of length  $N = 2^L$ , where  $L$  is an integer number. The wavelet representation of  $\mathbf{x}$  is given by:

$$\mathbf{x} = \nu_0 \boldsymbol{\nu} + \sum_{i=0}^{L-1} \sum_{j=0}^{2^i-1} w_{i,j} \psi_{i,j} \quad (3)$$

where  $\boldsymbol{\nu}$  is the scaling function and  $\psi_{i,j}$  is the wavelet function at scale  $i$  and offset  $j$ . Recalling our earlier matrix notation  $\mathbf{x}$  has an representation of  $\mathbf{x} = \Psi\boldsymbol{\alpha}$ , where  $\Psi$  is the Wavelet transform matrix and its columns contains the scaling and wavelet functions and  $\boldsymbol{\alpha} = [\nu_0, w_{0,0}, w_{1,0}, w_{1,1}, w_{2,0} \dots]^T$  is the vector of scaling and wavelet coefficients. The nested support of the wavelet at different scales create a parent/child relationship between wavelet coefficients at different scales. We say that  $w_{i,j}$  is the parent of its two children  $w_{i+1,2j}$  and  $w_{i+1,2j+1}$ . Considering sparsity for vector  $\boldsymbol{\alpha}$ , set of nonzero coefficients ( $\Omega = \text{supp}\{\boldsymbol{\alpha}\}$ ) forms a connected subtree, where if a coefficients  $w_{i,j} \in \Omega$  then its parent is in  $\Omega$  set as well. This motivates a connected tree model for the wavelet coefficients which is exploited in several works in reconstruction of smooth and piecewise smooth signals [10]. In a recent work [5], the same structure is used in a CoSaMP [11] greedy algorithm and Iterative Hard Thresholding (IHT) algorithm [12]. They proved that the required number of mea-

measurements  $M = \mathcal{O}(S)$  for a robust recovery of tree-sparse or tree-compressible signal.

### C. Enhanced Iterative Hard Thresholding (EIHT)

To overcome the aforementioned problems of the IHT approach, we propose to enhance the IHT recovery algorithm by adding a verification step after recursion step:  $\alpha_i = \mathcal{M}_k(\alpha_i + \Phi^T(\mathbf{y} - \Phi\alpha_i))$ . In each iteration, after hard-thresholding (or coefficients selection based on a model), the support vector of the coefficients in step  $i$  is  $\Omega_{S_i}$ :

$$\alpha_i := \{\alpha : \alpha|_{\Omega_{S_i}} \in \mathbb{R}^S, \alpha|_{\Omega_{S_i}^c} = 0\} \quad (4)$$

where  $\Omega^c$  denotes the complement of the set  $\Omega$ . In each iteration new candidates for the support set of  $\Omega$  are called  $\Omega_i^{new}$ :

$$\Omega_i^{new} : \{\Omega_{S_i} \cup \Omega_{S_{i-1}}\} - \{\Omega_{S_i} \cap \Omega_{S_{i-1}}\} \quad (5)$$

and each new candidate  $k \in \Omega_i^{new}$  is verified before adding to the support vector set:

$$\alpha_i = \begin{cases} \alpha_i|_{\Omega_i^{new}} = 0, \\ \forall k \in \Omega_i^{new} \nabla(f_{\Omega'_{S_i,k}}(\alpha'_{i,k}) - f_{\Omega_{S_{i-1}}}(\alpha_{i-1})) > 0 \end{cases} \quad (6)$$

where  $\Omega'_{S_i,k} = \Omega_{S_{i-1}} \cup k$ ,  $\alpha'_{i,k} := \{\alpha : \alpha|_{\Omega_{S_{i-1}} \cup k}\}$  and  $f(\alpha) = \|\mathbf{y} - \Phi\alpha\|$ . Step 6 ensures that in each iteration the new added values are in the right direction of the decreasing gradient of error. From an algorithmic point, when non of the new support candidate coefficients  $\Omega_i^{new}$  is in the correct path of decreasing gradient, we tune the values of  $\alpha$  by solving a Least Mean Squares problem according to:  $\alpha_j = \alpha_j - \mu \nabla f(\alpha)$ . This step will update the predicted values and continues the recursion.

Note that the enhancement proposed to IHT is also applicable for ALPS recovery algorithms. This is done in Section IV. The resulting recovery algorithms are indicated by a prefix "E". Additionally when the wavelet connected tree model is used in the recovery algorithm, the prefix "MB" is added.

## IV. RESULTS

To validate the performance of the different algorithms in recovering the compressed ECG signals, the MIT-BIH Arrhythmia Database [13] is used that is the most commonly used database for the comparative study of ECG compression algorithms. This database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects studied by the BIH Arrhythmia Laboratory. The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 10 mV range.

Moreover, to quantify the compression performance while assessing the diagnostic quality of the compressed ECG records, the *percentage root-mean-square difference (PRD)*, quantifies the percent error between the original signal vector  $\mathbf{x}$  and the reconstructed  $\tilde{\mathbf{x}}$ :

$$PRD = \frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2} \times 100, \quad (7)$$

Different values of PRD are classified based on the signal quality perceived by a specialist [14]. According to their investigation PRD values below than 2% and 9% are classified in "very good" and "good" reconstruction quality classes respectively.

In this section we present the results for different algorithms and we compare their performance in recovering the  $S$ -sparse and compressible signals. In all experiments ECG vector  $\mathbf{x}$  is a window of size  $N = 512$ . A random Gaussian matrix is used as our sensing matrix  $\Phi$ . The results are averaged over 100 packets of ECG data randomly selected from all database records. To apply the wavelet tree model in recovery algorithms, here we are using model-based compressive sensing toolbox [15], which uses the *condensing sort and select algorithm* (CSSA), a tree based approximation algorithm for finding the set of supports for best matching connected tree. Figure 1(a) shows the averaged output PRD for FLIHT, NIHT and EIHT for  $S$ -sparse signals with and without applying the wavelet connected tree model. For generating the  $S$ -sparse ECG vectors, set of supports for best connected tree including  $S$  coefficients are selected for each ECG packet, and all other coefficients are set to zero ( $S = 64$ ). Figure 1(b) also shows the probability of recovery for "good" reconstruction quality (defined as successful recoveries). As it is shown in figure 1(a) and 1(b), among all three recovery algorithms, EIHT shows the best performance for  $S$ -Sparse signals. But considering the model-based versions of the recovery algorithms, FLIHT and NIHT show a better recovery performance in the sense of the output reconstructed PRD and the probability of reconstruction. Figure 1(b) shows that model-based recovery MB-NIHT achieves more than 95% of successful recoveries for number of measurements  $M = 160 = 2.5 S$ . This is possible for  $M = 192 = 3 S$  for MB-FLIHT and  $M = 224 = 3.5 S$  for EIHT. Without applying a model this is only possible for  $M = 224 = 3.5 S$  for all recovery algorithms.

Similarly figure 2(a) does the same for FLIHT, EFLIHT, NIHT, ENIHT and EIHT for real ECG signals. The results show that EIHT shows superior performance, and then EFLIHT shows slightly better performance compared to the others. Then stands FLIHT and ENIHT and NIHT with very close performance. After applying the wavelet connected tree model, largest improvement in the performance results is for FLIHT and EFLIHT which shows a very close performance to the EIHT recovery algorithm.

Alternatively figure 2(b) compares the probability of recovery for "good" reconstruction quality as well. We observe that model based recovery "MB-FLIHT" and "MB-EFLIHT" achieves "good" quality signal recovery with probability over 90% with only  $M = 288$  number of measurements, while without applying a model they can only achieve this with  $M = 384$ .

## V. CONCLUSION

In this work, we proposed a new algorithmic improvement for sparse recovery algorithms based on support vector selection to reach a robust and stable recovery for compressible

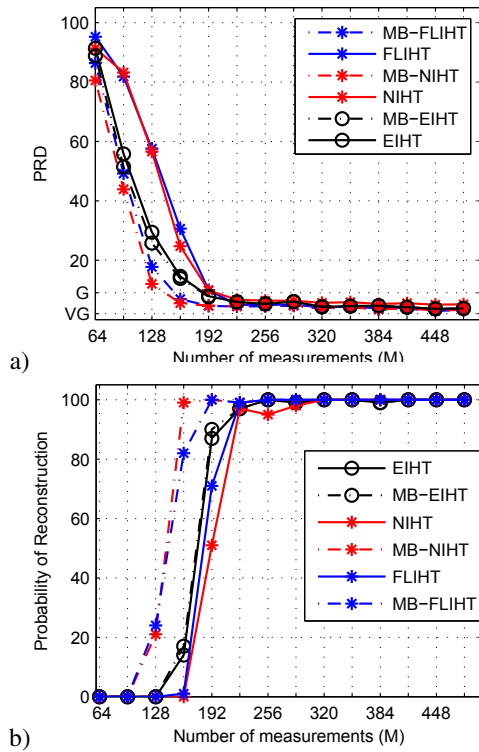


Fig. 1. Results for FLIHT, NIHT, and EIHT with and without applying model (prefix MB represent the model-based version), a) Output PRD averaged over all records for different number of measurements  $M$  for different algorithms. The records are forced to be exact  $S$ -sparse ( $S = 64$ ), b) Probability of reconstruction with "good" quality vs. number of measurements  $M$ .

signals. Moreover, for ECG signals, we showed that significant performance gains can be obtained by exploiting additional structural information of the underlying signal, beyond the simplistic sparse model. In particular, our results show that for real compressible ECG signals, the number of measurements can be decreased by 25% using a model-based recovery technique. However, despite the improved performance in terms of compression and robustness, the obtained results in the high-fidelity region are still not satisfactory. This suggests that the used model is not sufficient for describing the low-magnitude, yet, diagnostic-relevant details of ECG signals. Accordingly, more complex models for natural compressible -but not strictly sparse- biosignals may be investigated which take into account the well-known morphological dynamics of ECG.

## REFERENCES

- [1] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vandergheynst, "Compressed sensing for real-time energy-efficient ecg compression on wireless body sensor nodes," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 9, pp. 2456–2466, sept. 2011.
- [2] D. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, april 2006.
- [3] E. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, pp. 1207–1223, 2006.
- [4] R. G. Baraniuk, M. A. Davenport, R. A. DeVore, and M. B. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

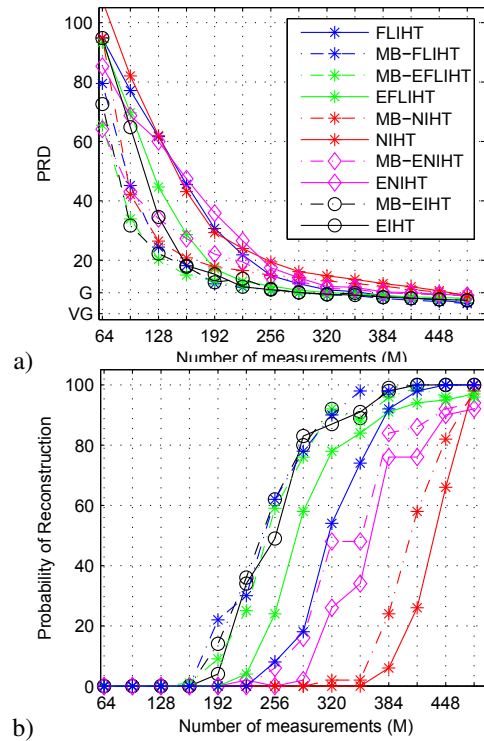


Fig. 2. Results for FLIHT, EFLIHT, NIHT, ENIHT and EIHT with and without applying model (prefix MB represent the model-based version), a) Output PRD averaged over all real ECG records for different number of measurements  $M$  for different algorithms, b) Probability of reconstruction with "good" quality vs. number of measurements  $M$ .

- [5] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, "Model-based compressive sensing," *Information Theory, IEEE Transactions on*, vol. 56, no. 4, pp. 1982–2001, april 2010.
- [6] V. Cevher, "An alps view of sparse recovery," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, may 2011, pp. 5808–5811.
- [7] Y. E. Nesterov, "A method for solving a convex programming problem with convergence rate  $\mathcal{O}(1/k^2)$ ," *Soviet Mathematics Doklady*, vol. 27, pp. 372–376, 1983.
- [8] V. Cevher, M. F. Duarte, C. Hedge, and R. Baraniuk, "Sparse signal recovery using Markov random fields," in *Neural Information Processing Systems (NIPS)*, 2008.
- [9] Y. Lu and M. Do, "Sampling signals from a union of subspaces," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 41–47, march 2008.
- [10] M. F. Duarte, M. B. Wakin, and R. G. Baraniuk, "Fast reconstruction of piecewise smooth signals from incoherent projections," in *Signal Processing with Adaptive Sparse Structured Representations (SPARS'05)*, November 2005.
- [11] D. Needell and J. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.
- [12] I. Daubechies, M. Defrise, and C. D. Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, pp. 1413–1457, 2004.
- [13] G. Moody and R. Mark, "The impact of the mit-bih arrhythmia database," *Engineering in Medicine and Biology Magazine, IEEE*, vol. 20, no. 3, pp. 45–50, may-june 2001.
- [14] Y. Zigel, A. Cohen, and A. Katz, "The weighted diagnostic distortion (wdd) measure for ecg signal compression," *Biomedical Engineering, IEEE Transactions on*, vol. 47, no. 11, pp. 1422–1430, nov. 2000.
- [15] [Online]. Available: <http://dsp.rice.edu/software/model-based-compressive-sensing-toolbox>