# Model Based Compressed Sensing Reconstruction Algorithms for ECG Telemonitoring in WBANs

Aris S. Lalos[a,*], Luis Alonso[a], Christos Verikoukis[b]

[a] *Department of Signal Theory and Communications (TSC) of the Technical University of Catalonia (UPC), Barcelona, Spain*
[b] *Telecommunications Technological Centre of Catalonia (CTTC), Castelldefels, Barcelona, Spain*

**Abstract**

Wireless Body area networks (WBANs) consist of sensors that continuously monitor and transmit real time vital signals to a nearby coordinator and then to a remote terminal via the Internet. One of the most important signals for monitoring in WBANs is the electrocardiography (ECG) signal. The design of an accurate and energy efficient ECG telemonitoring system can be achieved by: i) reducing the amount of data that should be transmitted ii) minimizing the computational operations executed at any transmitter/receiver in a WBAN. To this end, compressed sensing (CS) approaches can offer a viable solution. In this paper, we propose two novel CS based ECG reconstruction algorithms that minimize the samples that are required to be transmitted for an accurate reconstruction, by exploiting the block structure of the ECG in the time domain (TD) and in an uncorrelated domain (UD). The proposed schemes require the solutions of second - order cone

*Corresponding author. Tel: +34 93 413 71 47
E-mail Addresses : aristeidis.lalos@tsc.upc.edu (A. Lalos), luisg@tsc.upc.edu (L. Alonso), cveri@cttc.es (C. Verikoukis)

programming (SOCP) problems that are usually tackled by computational demanding interior point (IP) methods. To solve these problems efficiently, we develop a path-wise coordinate descent based scheme. The reconstruction accuracy is evaluated by the percentage root-mean-square difference (PRD) metric. A reconstructed signal is acceptable if and only if $PRD < 9\%$. Simulation studies carried out with real electrocardiographic (ECG) data, show that the proposed schemes, operating in both the TD and in the UD as compared to the conventional CS techniques, reduce the Compression Ratio (CR) by 20% and 44% respectively, offering at the same time significantly low computational complexity.

*Keywords:* Compressed Sensing, Block Sparsity, Wireless Body Area Networks, Real-time ECG monitoring.

---

## 1. Introduction

Recently, there has been increasing interest from researchers, system designers, and application developers on a new type of network architecture generally known as Wireless Body Area Network (WBAN). A WBAN is a collection of low-power, miniaturized, lightweight wireless sensor nodes that continuously monitor human's physiological activities and actions [1], such as health status and motion patterns. The real time vital signals are transmitted to a nearby body node coordinator (BNC) (e.g, smart phone) via ultra-low-power short-haul radios (e.g., ZigBee, Bluetooth), and then to a remote terminal (e.g., a hospital) via the Internet [1, 2]. One of the most important signal for monitoring and analyzing in WBANs is the electrocardiography (ECG) signal. The characteristic parameters of an ECG signal

(i.e., heart beat rates, morphology, dynamic behaviors) can be used for diagnosis of heart diseases such as myocardial ischemia, arrhythmia and cardiac infarction. ECG telemonitoring systems, relieve patients from the need of visiting hospitals frequently and allows the continuous and ubiquitous monitoring of their ECG.

Apart from the obvious advantages, ECG telemonitoring requires new schemes and algorithms to be implemented in order to optimize i) the energy consumption and ii) the total hardware cost both in the transmitter and the receiver. Note that 73% of the total power consumption at a biosensor, is consumed to the RF power amplifier (PA) [3]. As a result, low energy consumption significantly increases the battery lifetime of the biosensor. In addition, the hardware cost reduction makes the telemonitoring system economically viable and more easily accepted by the individual customers. Both requirements motivated the design of new ECG compression/reconstruction schemes, with high compression ratio capabilities and reduced computational requirements at both the transmitter and the receiver. The main drawback of conventional ECG compression scheme [4, 7, 10, 8, 9], is the increased computational compression requirements.

The fact that the ECG signal consists of both high activity (P, QRS and T waves) and low activity periods (e.g, isolectric lines between beats) that are less important, motivated researchers to develop schemes that exploit this information either at the transmitter or at the receiver. To exploit this signal property at the transmitter, low - rate adaptive sampling schemes [11, 12, 13, 14] may be employed. Those schemes exploit ECG signal characteristics by using a circuit that executes an activity detection algorithm that

identifies the high frequency regions of the signal and adjust the sampling rate [13]. An alternative solution that reduces both the processing executed and the data that should be transmitted, is the exploitation of the ECG signal characteristics at the receiver, by applying the compressive sensing (CS) theory for signal compression/reconstruction [15, 16, 17, 18]. By applying CS to the problem of ECG telemonitoring, the compression complexity at the transmitter is significantly reduced [17, 16], since the compression is performed through a simple linear encoding operation.

The CS based reconstruction of the original ECG signal is performed after digitization of the signal at the receiver by exploiting the structure of the original signal in both the time and wavelet domain. The authors in [17] achieved high compression rates (CR) of ECG and EMG biosignal in the time domain. Though in order to achieve that, they apply in advance a dynamic thresholding technique that firstly tracks the DC level of the signal and then tunes the signal sparsity by setting to zero the small signal values. The main drawback of this method is the increased computational processing at the transmitter and the fact that, useful parts of the signal can be destroyed during this process. The authors in [16] exploited the sparse structure of the ECG in the wavelet domain by employing conventional compressed sensing (CS) schemes. However, the achieved compression ratio (CR) as compared to traditional discrete wavelet based ECG compression schemes [10], where a fixed percentage of wavelet coefficients are zeroed, was 20% lower. Nevertheless, the significantly low compression complexity renders the CS method, promising in wireless ECG telemonitoring applications. Finally, Zhang et al. [18] was the first that proposed the use of a generic block sparse Bayesian

learning framework to reconstruct fetal ECG recordings by exploiting the block sparsity. The authors showed that the aforementioned framework can reconstruct the fetal ECG recordings with high quality but with increased computational cost.

To the best of our knowledge, this is the first CS based work that exploits in a computational efficient way, key characteristics (e.g., block sparsity, intra-block correlation and presence of dc offset) of the adult ECG signal in both the time domain (TD) and in an uncorrelated domain (UD) that further promotes block sparsity. We propose efficient low complexity compression/recovery algorithms that are computational efficient and achieve significant CR reduction, compared to conventional CS approaches [17, 16]. The proposed schemes are the first CS based schemes that outperform even the computational demanding discrete wavelet based ECG compression schemes [10, 8, 9]. Our main contributions can be summarized in the following:

- We eliminate the need of estimating and removing the dc level of the ECG signal at the transmitter (i.e., biosensor) [17] by considering it as an unknown parameter that can be estimated at the receiver. This modification allows us also, to track any abrupt shift in baseline due to movement of the adult patient while the ECG is being recorded [20].

- We propose two novel ECG reconstruction schemes that exploit the block sparse structure of the transmitted signal in the TD and in an UD, that further promotes block sparsity, by formulating second-order cone programming (SOCP) problems which can be solved by a host of computational demanding interior point (IP) methods. These schemes achieve the same reconstruction accuracy with the default CS

approaches, by using 20% and 44% less samples respectively. In addition, we show through extensive evaluation that the proposed schemes achieve CRs that are even higher than the traditional discrete wavelet based ECG compression schemes [10, 8, 9], which are much more computational demanding.

- We develop a novel scheme that is based on the path-wise coordinate descent approach [21] and significantly reduce the computational cost of the conventional IP methods, applied for the recovery of the ECG. The derived low complexity schemes achieve similar recovery properties with those that apply IP methods, offering at the same time significant savings in complexity.

The paper is outlined as follows: Section 2, introduces the concepts and terminology related with CS theory and gives a short description of the structure of the ECG signal in TD. In Section 3, we present several low complexity encoding strategies with different requirements (i.e. communication, storage, computational) that may be employed in the biosensor, for compressing the ECG signal. Section 4, introduces novel recovery algorithms that exploit the structure of the ECG signal either in the TD or in an UD by employing computational demanding IP methods. To reduce the complexity of those methods, we propose a low complexity path-wise coordinate descent based solver. The proposed schemes are evaluated through extensive simulations which are presented in Section 5, followed by conclusions in Section 6.

## 2. Preliminaries

### *2.1. Compressive Sensing Framework*

CS provides approaches for reconstructing a sparse signal $\mathbf{x} \in \mathbb{R}^N$ by using a small number of linearly coded measurements [23, 24, 25]. The encoded measurements, $\mathbf{y} \in \mathbb{R}^M$, are generated using a random matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$ with independent and identically distributed (i.i.d.) elements. In mathematical terms, $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$, where $\mathbf{w}$ is a vector with noise samples.

#### *2.1.1. Reconstruction by Exploiting Sample Sparsity*

In the noise free case (i.e., $\mathbf{w} = \mathbf{0}_N$), the vector $\mathbf{x}$ may be ideally recovered from the vector $\mathbf{y}$ by solving the minimization problem $min_x \{\|\mathbf{x}\|_0 : \mathbf{y} = \mathbf{A}\mathbf{x}\}$. In order to make the signal reconstruction robust to the presence of noise, the constraint of the problem may be relaxed as $min_x \{\|\mathbf{x}\|_0 : \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \leq \epsilon\}$, where $\epsilon$ is a predefined error tolerance. This approach, however, cannot be used for practical applications, since it is computationally intractable. CS provides a solution to this issue, by replacing the $\ell_0$ quasi-norm with the convex $\ell_1$-norm and by solving a new optimization problem that can be formulated as $min_x \{\|\mathbf{x}\|_1 : \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \leq \epsilon\}$, where the $\ell_1$-norm is defined as $\|\mathbf{x}\|_1 = \sum_{i=1}^{N} |x_i|$. By employing Lagrange relaxation, it is possible to efficiently find an approximate solution by solving the problem:

$$\hat{\mathbf{x}} := \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1, \tag{1}$$

where the parameter $\lambda$ controls the balance between two optimization objectives: i) the noise level $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$ and ii) the sparsity of vector $\mathbf{x}$. Algorithmically, the convex optimization problem in eq. (1), known as LASSO problem, can be tackled by any Second-Order Cone Program (SOCP) solver.

*2.1.2. Reconstruction by Exploiting Block Sparsity*

A block sparse signal is a signal that consists of clusters of non-zero coefficients.In mathematical terms, a vector $\mathbf{x}$ can be formulated as a concatenation of $R$ blocks of length $d$, as follows:

$$\mathbf{x} = [\ \underbrace{x_1, \ldots, x_d}_{\mathbf{x}^T[1]}, \underbrace{x_{d+1}, \ldots, x_{2d}}_{\mathbf{x}^T[2]}, \ldots, \underbrace{x_{N-d+1}, \ldots, x_N}_{\mathbf{x}^T[R]}\ ]^T, \tag{2}$$

where $\mathbf{x}[i]$ denotes the $i^{th}$ block and $N = Rd$. Vector $\mathbf{x}$ is called block $k$-sparse if $\mathbf{x}[i]$ consists of $d$ zeros for at most $R - k$ indices $i$. Similar to eq. (2), the measurement matrix $\mathbf{A}$ can be represented as a concatenation of sub-matrices $\mathbf{A}[i]$ of size $M \times d$, i.e., $\mathbf{A} = [\ \mathbf{A}[1]\ \ \mathbf{A}[2]\ \ \ldots\ \ \mathbf{A}[R]\ ]$.

By using $\ell_1$-relaxation for reconstructing $\mathbf{x}$, we ignore the fact that the signal is block-sparse, i.e., the non-zero entries occur in consecutive positions. To exploit block sparsity, it is possible to reconstruct the vector $\mathbf{x}$ by solving the optimization problem as follows:

$$\hat{\mathbf{x}} := \arg\min_{\mathbf{x}} \|\mathbf{y} - \sum_{i=1}^{R} \mathbf{A}[i]\,\mathbf{x}[i]\,\|_2^2 + \sum_{i=1}^{R} \lambda \|\mathbf{x}[i]\,\|_2, \tag{3}$$

which is also known as group LASSO problem [21, 22].

*2.2. ECG Signal Model*

By inspecting fig. 1 it is easily shown that the location of the significant coefficient (i.e. the QRS complex, as well as T and P waves) cluster in blocks. We assume that the signal to be transmitted contains noise itself and a dc offset that may vary in time due to the movement of a patient while his ECG is being recorded [20][1]. Thus, the samples that correspond to the

[1]According to [20, fig. 6] an abrupt shift of the baseline may occur due to movement of the patient while the ECG is being recorded.

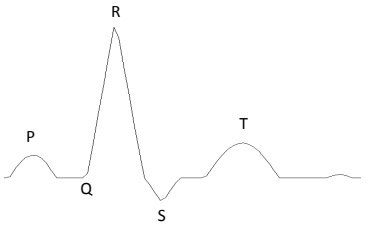



Figure 1: *Typical one cycle ECG tracing.*

low activity regions (e.g., isoelectic line), which contain less information are usually larger than zero due to the presence of the dc bias. Therefore, under the assumption that the *dc* value remains constant during the sampling of an $N$-sample segment, a recorded ECG segment may be written as:

$$\mathbf{u} = \mathbf{x} + d_c \mathbf{1}_N + \mathbf{w}, \tag{4}$$

where $\mathbf{u} = [u_1, \ldots, u_N]^T$ are samples of a noisy signal segment, $\mathbf{1}_N$ is an $N \times 1$ vector of all ones, $d_c$ is a scalar value that correspond to the unknown dc level of the signal, $\mathbf{x} = [x_1, \ldots, x_N]^T$ where $x_i \in \mathbb{R}$ is a block sparse signal, and $\mathbf{w} = [w_1, \ldots, w_N]^T$ is the signal noise. To be able to efficiently apply the conventional CS based methods presented in Section 2, we should remove this dc bias from the recorded signal values, prior to the encoding process that is described in Section 3. By making this formulation we are able to estimate both the $d_c$ value and the vector $\mathbf{x}$ coefficients at the receiver, eliminating the need of tracking it and removing $d_c$ at the transmitter [17].

## 3. Compression of the ECG Segment

For each segment we perform compression by generating $M < N$ random linear combinations as follows:

$$\mathbf{y} \;=\; \mathbf{Au} = \mathbf{Ax} + d_c\mathbf{A1}_N + \tilde{\mathbf{w}} \tag{5}$$

where $\tilde{\mathbf{w}} = \mathbf{Aw}$ and $\mathbf{A}$ is a matrix of dimension $M \times N$, with stochastically independent random entries. There are several ways of constructing matrix $\mathbf{A}$ that directly affects i) the storage and processing requirements at the transmitter side, ii) the communication load and iii) the storage requirements and the recovery error at the receiver. Below we present the most widely used sensing matrices in the literature of CS [28]:

*Gaussian Random Encoding*

By selecting the coefficient $\mathbf{A}_{i,j} \sim \mathcal{N}\left(0, 1/\sqrt{N}\right)$ as Gaussian i.i.d. elements the recovery conditions are satisfied. Even though the authors in [16] showed that the quantization of the Gaussian elements do not affect significantly the signal quality loss, the aforementioned choice requires i) the implementation of a Gaussian distributed random generator, ii) the multiplication of the ECG samples with real values iii) the storage of a large matrix with real values.

*Binary Random Encoding*

An alternative approach is to select the i.i.d. entries of matrix $\mathbf{A}$ from the Bernoulli distribution, i.e. $\mathbf{A}_{i,j} = \pm 1/\sqrt{N}$ with probability 0.5. It has been shown that the aforementioned matrix reduces significantly the processing

at the transmitter, while satisfying the recovery conditions. Note that the transmitted data are generated by simply adding/subtracting the original ECG samples.

*Sparse Binary Random Encoding*

To reduce even more the required compression complexity, the entries $\mathbf{A}_{i,j}$ may be selected according to [29]: $\mathbf{A}_{i,j}$ are either $\{1, 0\}$ with probabilities $\{1/s, 1 - 1/s\}$ where $s$ is a parameter that determines the degree of sparsity of the sensing matrix $\mathbf{A}$. The authors in [29] showed that sparse random matrices still yields good recovery properties. The optimal choice of $s$ depends on the structure of the ECG signal and the decoding algorithm that is used at the receiver side.

To increase even more the compression efficiency of the CS based encoder presented above, an architecture similar to the one presented in [16] may be adopted. Fig. 2 presents a block diagram of both the transmitter and receiver architecture. According to this diagram, the encoded samples $\mathbf{y}$ pass through: i) a redundancy removal unit ii) a quantization unit and iii) a Huffman encoding unit. According to [16], in cases where we perform compression with the same random matrices, a large inter-packet redundancy is introduced. This redundancy should be removed prior to Huffman encoding and wireless transmission. To that end, the redundancy removal unit computes the difference between consecutive vectors, and only this difference is further processed. It was shown in [16], that the output of the redundancy removal unit can be sufficiently represented using 9 bits. As a result, the quantization unit performs quantization using a 512 level uniform quan-

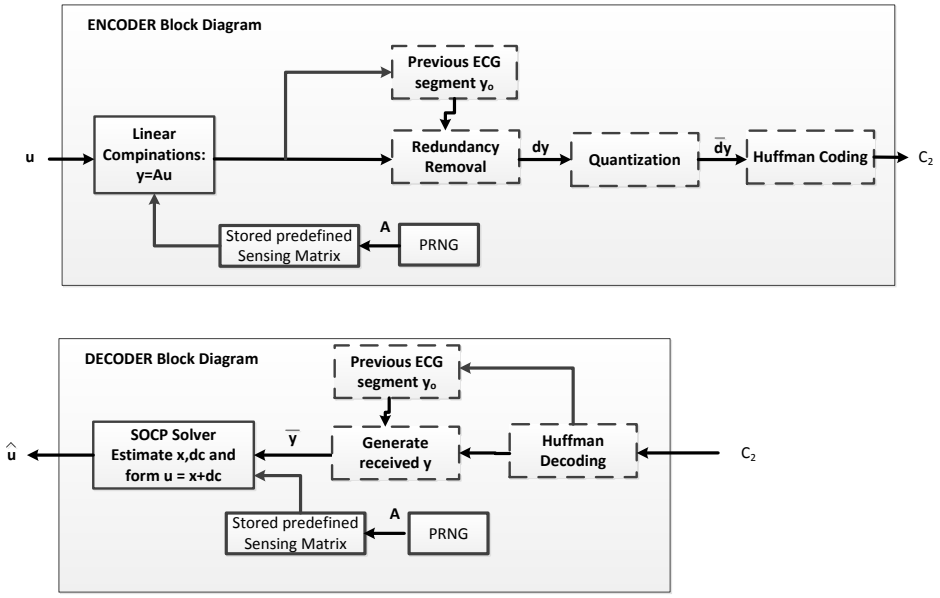



Figure 2: Block Diagram of the CS based Encoder/Decoder Architecture.

tizer.[2] Finally, the Huffman encoding unit performs loss-less Huffman coding to generate the binary encoding vector denoted by $C_2$.

In general, it is assumed that the encoding matrix $\mathbf{A}$ is known at the destination in order to perform reconstruction of $\mathbf{x}$,$d_c$ from $\mathbf{y}$. To overcome this limitation, a pseudo-random number generator (PRNG) that generates a sequence of numbers that approximates the properties of random numbers, is usually adopted at both the transmitter and the receiver. The generated sequence is completely determined by a relatively small set of initial val-

[2]Note that more efficient quantization schemes can lead to even higher CRs but the investigation of those schemes is beyond the scope of this study.

ues, called PRNG state, which includes a truly random seed that has to be transmitted at the receiver side. To reduce even more the communication requirements, a generated random matrix of dimensions $M \times N$ is stored and used for encoding a large number of ECG segments of size $N$. In the following section, we describe the operations executed on the receiver side, to recover the original ECG signal $\mathbf{u}$ from the encoded vector $\mathbf{y}$ given the encoding matrix $\mathbf{A}$.

## 4. Efficient Recovery of the original ECG Signal

A simplified block diagram of the decoder architecture is also provided in fig. 2. In the receiver, given the encoded vector $\mathbf{y}$ and matrix $\mathbf{A}$, an efficient CS based recovery procedure that exploit the structure of the ECG signal is taking place. This section presents two proposed schemes that exploit this structure in both the TD and in an UD. The first one, initially identifies and removes the dc level of the specific segment. Then after discarding the part (blocks) of the signal that contain less information in the TD, it estimates the parts that are likely to correspond to the high activity regions (e.g., P,Q,R,S and T waves). The second scheme can then be directly derived from the first one after further taking into account the amplitude correlation among the elements within each block (intra-block correlation). In that way, we significantly reduce the required samples at the receiver that are essential for achieving a relatively small reconstruction error. The identification of the dc level and the estimation of the ECG blocks, are written as SOCPs that can be solved by using classical interior point (IP) methods [30]. To reduce the computational complexity of those methods, we present a scheme

that is based on the path-wise coordinate descent approach and achieves performance almost identical to that of the IP methods at a much lower operational cost.

*4.1. Exploitation of the Block Sparse Structure in the TD*

To identify the clusters of samples that correspond to parts of the P,Q,R,S and T waves, we need to reformulate problem (3) as follows:

$$\min_{\mathbf{x},d_c} \|\mathbf{y} - \sum_{i=1}^{R} \mathbf{A}\,[i]\,\mathbf{x}\,[i] - d_c\mathbf{A}\mathbf{1}_N\|_2^2 + \sum_{i=1}^{R} \lambda\|\mathbf{x}\,[i]\,\|_2 \tag{6}$$

where the dc offset is considered an unknown parameter of the specific ECG segment. This allows us to exploit even more the structure of each segment itself and estimate the unknown dc level at the receiver. Given vector $\mathbf{x}$, the least squares (LS) estimation of the unknown dc offset is given by:

$$\hat{d}_c = \left(\mathbf{a}^T\mathbf{a}\right)^{-1}\mathbf{a}^T\left(\mathbf{y} - \mathbf{A}\mathbf{x}\right) \tag{7}$$

where $\mathbf{a} = \mathbf{A}\mathbf{1}_N$. By substituting eq. (7) in (6) we can rewrite the aforementioned problem as:

$$\min_{\mathbf{x}} \|\mathbf{y}_n - \mathbf{A}_n\mathbf{x}\|_2^2 + \sum_{i=1}^{R} \lambda\|\mathbf{x}\,[i]\,\|_2 \tag{8}$$

$$\mathbf{y}_n = \left(I_M - \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}\right)\mathbf{y}, \mathbf{A}_n = \left(I_M - \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}\right)\mathbf{A}. \tag{9}$$

The aforementioned group LASSO problem can be solved by using standard software packages, that employ IP methods. After evaluating the solution of the problem in (8), we are able to evaluate the original ECG segment $\mathbf{x}$ having completely removed the dc offset coefficient, which in turn can be

estimated by eq. (7). The estimation accuracy of the aforementioned procedure, can be improved by using an iterative reweighed approach. This can be intuitively justified by the following remark:

**Remark 1.** *Using an iterative algorithm to construct the weights ($w_i$) allows us a better estimation of the blocks with the nonzero coefficient. Although, the early iteration execution may lead to inaccurate signal estimates, the blocks with largest signal coefficients are most likely to be identified as nonzero. Once these blocks are identified, their influence is down-weighted in order to increase sensitivity during the identification of the remaining blocks with small but nonzero signal coefficients. Thus, the recorded measurements* $\mathbf{u}$ *may be estimated more accurately, by executing the iterations:*

$$\mathbf{x}^{(l)} := \arg\min_{x} \|\mathbf{y}_n - \mathbf{A}_n\mathbf{x}\|_2^2 + \sum_{i=1}^{R} \lambda w_i^{(l)} \|\mathbf{x}\,[i]\,\|_2 \tag{10}$$

$$w_i^{(l)} := \left(\|\mathbf{x}^{(l-1)}\,[i]\,\|_2 + \epsilon\right)^{-1}, \quad i = 1, \ldots, R \tag{11}$$

*Proof.* The analytical justification is based on the connection of the reweighed method with the log sum penalty, since it has been shown, that the log sum function is much more sparsity encouraging than the $l_1$ norm [31]. To be more specific, it can be derived form the justification given in [31, Section 2.3] by substituting the $l_1$ norm with the term $\sum_{i=1}^{R} \|\mathbf{x}\,[i]\,\|_2$. □

The optimization per iteration of (10) is a weighted version of (8) and thus can be efficiently solved in a similar manner as the problem defined in (8). The iterations can be initialized with the (8) solution which corresponds to setting all weights to unity. The iterative scheme can be terminated as soon

as the relative error $\|\mathbf{x}^{(l)}-\mathbf{x}^{(l-1)}\|_2/\|\mathbf{x}^{(l)}\|_2$ becomes smaller than some chosen $\epsilon$ equal to say $10^{-6}$. The aforementioned iterative re-weighted method further improves the accuracy of both the estimated non zero blocks and the variable dc offset. The benefits of the proposed algorithm in terms of convergence (e.g., required samples for achieving accurate reconstruction at the receiver), are evaluated in section V.

*4.2. Exploitation of the Block Structure of the Uncorrelated ECG*

In this subsection we provide modifications that can be applied in the problem defined by eqs. (10), (11) in order to efficiently exploit the block sparsity of the ECG signal in an UD. It is shown that the eliminations of any possible intra-block correlation (i.e., the amplitude correlation among the elements within each block) can further promote block sparsity and therefore improve the reconstruction performance of the aforementioned schemes.

Initially, let us assume that i) a Toeplitz matrix $\mathbf{T}_i \in \Re^{d\times d}$ captures the correlation structure of the $i$-th block and ii) the correlation between elements of different blocks is zero, i.e.,

$$E\left[\mathbf{x}\left[i\right]\mathbf{x}^T\left[j\right]\right] = \begin{cases} \mathbf{T}_i & if\ i=j \\ 0 & if\ i\neq j \end{cases}, \tag{12}$$

$$\mathbf{T}_i = \begin{bmatrix} r_0 & r_1 & \dots & r_{d-1} \\ \vdots & \ddots & \ddots & \vdots \\ r_{d-1} & \dots & r_1 & r_0 \end{bmatrix} \tag{13}$$

The value of $r_k$, $\forall k = 0,\dots,d-1$ can be calculated from the ECG samples by using an exponentially decaying data window [40]. Alternatively, we can

make the approximation $r_k = (r)^k$ and select only a fixed value for $r$ that captures the correlation between adjacent samples (i.e., $r = 0.94$). To exploit the block sparsity of the ECG signal in an UD we need to apply the following transformation matrix $\mathbf{T}^{-1/2}$ to the original ECG segment:

$$\mathbf{s} = \mathbf{T}^{-1/2}\mathbf{x}, \tag{14}$$

$$\mathbf{T}^{-1/2} = \begin{bmatrix} \mathbf{T}_1^{-1/2} & \mathbf{0}_d & \dots & \mathbf{0}_d \\ \dots & \ddots & \ddots & \dots \\ \mathbf{0}_d & \dots & \dots & \mathbf{T}_R^{-1/2} \end{bmatrix}, \tag{15}$$

where $\mathbf{s}$ denotes the transformed signal and $\mathbf{0}_d$ is a $d \times d$ matrix of all zeros. This procedure further promotes block sparsity since as it will be shown in the simulation section the samples of the prewhittened signal vector $\mathbf{s}$ can be efficiently recovered by less encoded samples than the samples of the original signal vector $\mathbf{x}$. This remark can be intuitively justified by the fact that group lasso algorithms become more efficient or equivalently provide better reconstruction accuracy for a given number of encoded samples, when the difference between the norms of non-zero blocks is minimized [26]. In fig. 3 we provide the original and decorrelated version of an ECG segment that correspond to a 2 sec recording of a male ECG, taken from the MIT-BIH Normal Sinus Rhythm Database [34], where it can be seen that this property is more valid for the prewhittened signal than the original one.

To recover the original vector at the receiver by controlling the block sparsity of the the prewhittened ECG segment $\mathbf{s} = \mathbf{T}^{-1/2}\mathbf{x}$ we can apply the algorithm presented in eqs. (10), (11) after replacing the $\ell_2$ norm of $\mathbf{x}[i]$ with $\|\mathbf{s}[i]\|_2$. More specifically, in the $l$-th iteration we need to solve the following optimization problem:

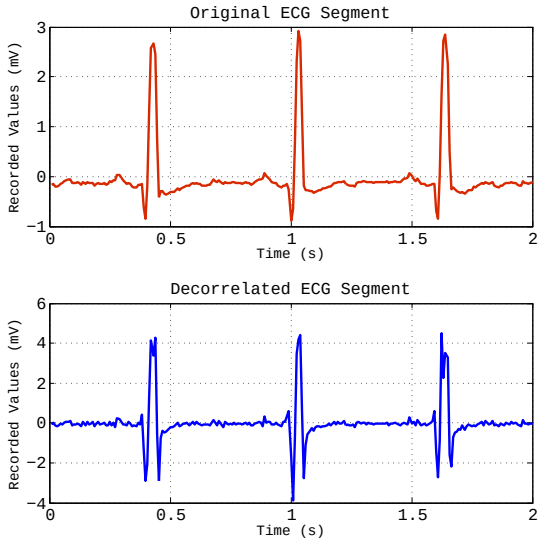



Figure 3: *Original and decorrelated ECG segment values by applying eq. (14) with a fixed* $r = 0.94$

$$\mathbf{s}^{(l)} := \arg\min_{s} \|\mathbf{y}_n - \mathbf{A}_n \mathbf{T}^{1/2} \mathbf{s}\|_2^2 + \sum_{i=1}^{R} \lambda w_i^{(l)} \|\mathbf{s}\,[i]\,\|_2 \tag{16}$$

$$w_i^{(l)} := \left(\|\mathbf{s}\,[i]\,\|_2^{(l-1)} + \epsilon\right)^{-1}, \quad i = 1, \ldots, R \tag{17}$$

where matrix $\mathbf{T}_i$ may be computed according to eq. (13). To reduce the number of iterations that need to be executed by the proposed algorithm, we select to predefine the value of $r_k^{(l)} = (r)^k$ and set $r = 0.94$. Each iteration $l$, defined in eq. (16) is a group LASSO iteration and the ECG blocks samples are computed by $\mathbf{x}^{(l)}\,[i] = \mathbf{T}_i^{1/2} \mathbf{s}^{(l)}\,[i]$. The steps of the proposed algorithm that further exploits the intra-block correlation between samples of the same block are summarized in Table 1.

*4.3. Block Coordinate Descent: An Efficient Group LASSO Solver*

Algorithmically, the problems in eqs. (10) and (16) are tackled by IP schemes customized to its specific form [35]. The main drawback of those schemes is the increased complexity, which is $O\left(M^2N^{1.5}\right)$. To reduce the computational cost of an IP solver of (10), (16) by $O\left(\sqrt{N}\right)$, we follow the procedure of the coordinate descent scheme presented in [21]. For the sake of simplicity, we focus on how the algorithm can be applied in one iteration of (10) and thus we drop index notation $l$ in $\mathbf{x}^{(l)}[i]$,$w_i^{(l)}$. The core idea is to iteratively minimize eq. (3) w.r.t. one block $\mathbf{x}[i]$ at a time, while keeping the remaining ones fixed:

$$\min_{\mathbf{x}[i]} \|\mathbf{y}_n - \sum_{j\neq i} \mathbf{A}_n[j]\,\mathbf{x}[j] - \mathbf{A}_n[i]\,\mathbf{x}[i]\,\|_2^2 + \lambda w_i \|\mathbf{x}[i]\,\|_2. \tag{18}$$

Then we repeat the minimization in a round robin way, for a given number of iterations. According to [21], the aforementioned problem admits a closed form solution that can be written as:

$$\hat{\mathbf{x}}[i] = \mathbf{\Phi}_{A_n[i]}^{-1} \frac{\mathbf{g}[i]}{\|\mathbf{g}[i]\,\|_2} \max\left(\|\mathbf{g}[i]\,\|_2 - \lambda w_i, 0\right) \tag{19}$$

where $\mathbf{\Phi}_{A_n[i]} = \mathbf{A}_n^H[i]\,\mathbf{A}_n[i] + \delta\mathbf{I}_{d_i}$. The proposed block coordinate descent algorithm (BCD) iterates by applying eqs. (7), (19) until a certain condition is met (e.g., a given number of iteration, or a given error tolerance). Since $\mathbf{g}[i] = \mathbf{A}_n^T[i]\left(\mathbf{y}_n - \sum_{j\neq i}\mathbf{A}_n[j]\,\mathbf{x}[j]\right)$, vector $\mathbf{g}[i]$ can be updated as $\mathbf{g}[i] = \mathbf{g}[i] + \mathbf{A}_n[i]\,\hat{\mathbf{x}}[i]$. After updating the value of $\hat{\mathbf{x}}[i]$ the value of $\mathbf{g}[i]$ is updated as $\mathbf{g}[i] = \mathbf{g}[i] - \mathbf{\Phi}_{A_n[i]}\hat{\mathbf{x}}[i]$. Note that since the matrices $\mathbf{\Phi}_{A_n[i]}^{-1}$ may be computed and stored offline, the most computationally demanding operation is the matrix-vector products $\mathbf{\Phi}_{A_n[i]}^{-1}\mathbf{g}[i]$, $\mathbf{\Phi}_{A_n[i]}\mathbf{x}[i]$.

## 5. Performance Evaluation

The focus of this study is to identify the benefits of exploiting different key characteristics during reconstruction of the ECG signal in a WBAN receiver. In this section, we present the simulation setup along with the results of our experiments.

### *5.1. ECG Dataset and Performance metrics*

For the evaluation of the proposed schemes we use both the MIT-BIH normal sinus rhythm database and the MIT-BIH compression database [34], where type of arrythmias are also present. The MIT-BIH Normal Sinus Rhythm database includes 18 long-term ECG recordings of subjects referred to the Arrhythmia Laboratory at Boston's Beth Israel Hospital (now the Beth Israel Deaconess Medical Center) from 5 men, aged 26 to 45, and 13 women, aged 20 to 50. The recordings were digitized at 128 samples per second per channel. While the MIT-BIH compression database contains 168 short ECG recordings (20.48 seconds each) selected to pose a variety of challenges for ECG compressors, in particular for lossy compression methods. The recordings in this database were digitized at 250 samples per second per channel at 12 bit resolution.

We assume that each ECG signal is divided into segments of $N = 256$ samples in the case of MIT-BIH normal sinus database and $N = 512$ samples in the case of MIT-BIH compression database, in order to refer to similar signal segments since the signal in the two different cases is digitized with different sampling rates. To simulate an abrupt shift to the dc level due to movement of the patient during recording of the ECG [20, fig. 6], we

introduce a variable dc offset in each ECG segment, the value of which is selected randomly in the interval $[0 - 100]$. Each segment $\mathbf{y}$ is encoded by using any of the matrices of section 3, of dimensions $M \times 256$. To perform a more complete and realistic study, we study both cases where either the three units that perform inter-packet redundancy removal, quantization and Huffman encoding are considered or not (see, fig. 2).

In the receiver side, given a specific number of received samples $M$, after the Huffman decoding procedure (whenever the aforementioned units are considered), we perform decoding of the original $N$ samples by using the following CS based estimators: i) the classical LASSO approach that uses a standard SOCP solver (interior point (IP) scheme) of (6) (LASSO - IP), ii) the Cyclic Coordinate Descent (CCD) approach that occurs from the BCD algorithm of Section 4.3 for $d = 1$ iii) the iterative reweighed group LASSO (IG LASSO-IP) of Table 1 that uses IP methods for solving iteratively the problem defined in eq. (10) iv) the IG LASSO of Table 1 that uses the BCD scheme (IG LASSO - BCD) for approximating the solutions of (10) v) the prewhittened IG LASSO (PWIG LASSO - IP) of Table 1 that uses IP methods for solving (16) vi) the PWIG LASSO that employs BCD scheme (PWIG LASSO - BCD) for solving (16). The scaling rules for the parameter $\lambda$ in the problems (10) and (16) follow the results of [37] and [38].

Finally, we compare the CS based compression/ reconstruction schemes with the traditional discrete wavelet based ECG compression scheme, that consider a thresholding-based algorithm presented in [10] (estimator (vii)), where a fixed percentage of wavelet coefficients are zeroed. This scheme, so called as (TH DWT) has been selected as the baseline scheme since it has

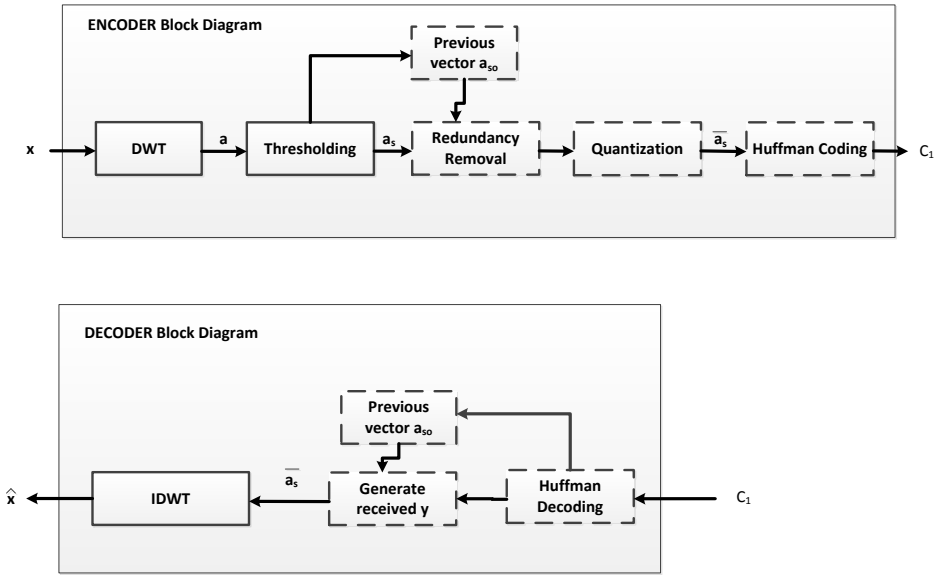



Figure 4: Block Diagram of the TH DWT Encoder/Decoder.

been shown to outperform many discrete wavelet transform (DWT) based ECG compression algorithms such as: i) the embedded zero tree wavelet [8], ii) the set partitioning in hierarchical trees [9]. Note that although the authors in [10] used the bi-orthogonal bior4.4 wavelet, in this study we utilize the orthogonal Daubechies wavelets (db 10) as the most popular wavelet family for ECG compression [16]. The block diagram of the baseline (TH DWT) encoder and decoder is presented in fig. 4. The DWT of the original signal is computed and then the $N - M$ smallest coefficients are zeroed. To select the $N - M$ smallest coefficients, a special unit that computes a dynamic threshold that is strongly connected to the value of $M$ should be added. Each ECG coefficient is then compared to this threshold and if its value is lower than the threshold is set to zero.

The diagnostic quality of the compressed ECG recordings, is evaluated by using the percentage root-mean-square difference (PRD) [36] that is defined as: $PRD = \|\mathbf{u} - \tilde{\mathbf{u}}\|_2 / \|\mathbf{u}\|_2 \times 100$, where $\tilde{\mathbf{u}} = \tilde{\mathbf{x}} + \tilde{dc}\mathbf{1}_N$ is constructed from the estimated value $\tilde{dc}$ and vector $\tilde{\mathbf{x}}$. In our scenarios, we assume that a reconstructed signal is acceptable if and only if $PRD < 9\%$ [36]. The aforementioned algorithms are evaluated either in terms of the average PRD, or in terms of the success rate, defined as the following probability $Pr\{PRD < 9\%\}$. Finally, the compression quality of the transmitted ECG, is evaluated by using the compression ratio (CR). The CR is either calculated at the output of the random linear coding unit in fig. 2 (or the DWT unit in fig. 4) by $CR = (N - M)/N \times 100$, or at the output of the Huffman coding unit (in both figs. 2, 4) by $CR = \frac{12 \times N - L_2}{12 \times N} \times 100$. With $L_2$ we denote the length of the encoding vector $C_2$ and $12 \times N$ are the total bits required for the representation of the $N$ initial recorded ECG samples, using a 12 bit resolution.

*5.2. Performance Results*

In order to identify the more efficient CS based encoding schemes that achieve the optimal trade off between computational requirements at the transmitter and reconstruction error at the receiver, we conducted experiments that employ the different encoding strategies presented in section III. The CR ratio is computed at the output of the random linear encoder. Decoding is performed by employing estimator (iii) and is evaluated by the estimated success rate. The block length was set equal to $d = 16$. Fig. 5, it can be easily shown that the use of a sparse binary matrix and sparsity parameter equal to $s = 4$, results in a success rate similar to that achieved

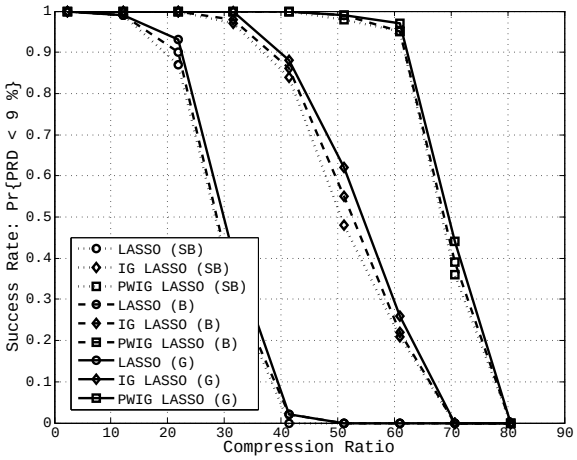



Figure 5: *Evaluation of different encoding methods (e.g, Gaussian (G), Binary (B) and Sparse Binary (SB) with s:4) in terms of PRD. For solving problems (10) and (16) we made use of IP methods. The CR has been calculated at the output of the random linear encoder. Database: MIT BIH normal sinus*

when using a Gaussian encoding matrix, while at the same time eliminates completely the need of multiplications at the transmitter[3] and reduces the number of additions/subtractions by 50%. Therefore for the rest of our study we make use of a sparse binary encoding matrix with $s = 4$.

To efficiently select the appropriate block length $d$ for the block CS based algorithms (e.g., IG LASSO and PWIG LASSO), we studied the effects of group size $d$ in terms of the success rate. Again the CR is calculated at the output of the linear encoder. The results are presented in fig. 6, where it is shown that as $d$ increases, the performance of the PWIG LASSO scheme

[3]Normalization with the $\sqrt{N}$, is implemented in Hardware, by performing a simple shifting operation.

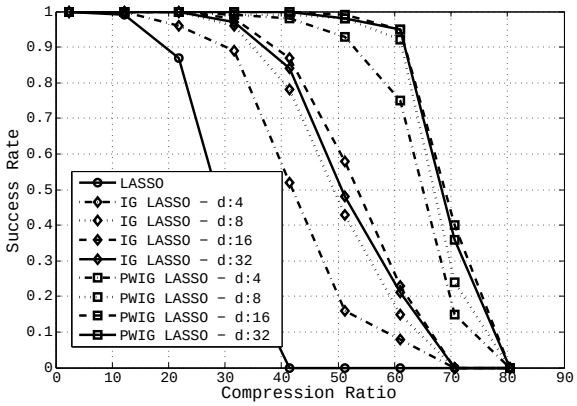
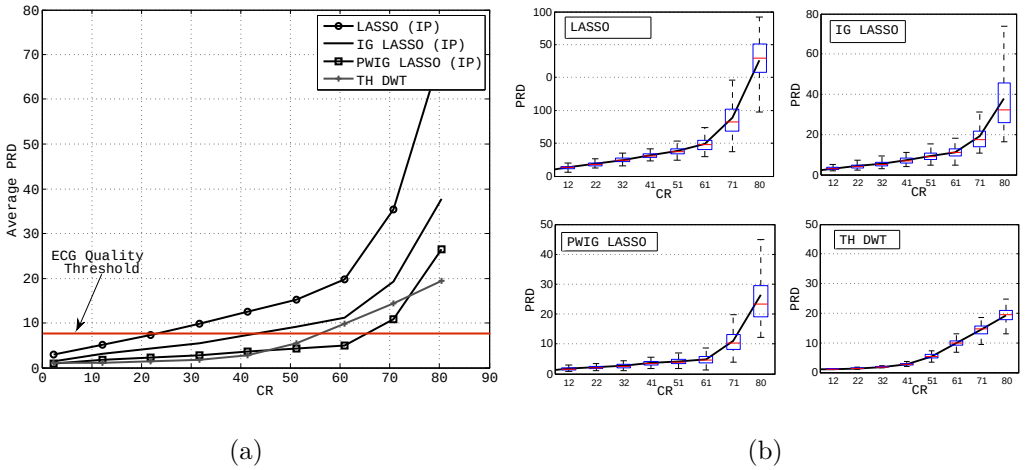



Figure 6: Success rate: ($Pr\{PRD < 9\%\}$) at the output of (a) the random linear encoder (b) the huffman encoder. Database: MIT BIH normal sinus.




(a) (b)

Figure 7: *Evaluation of the reconstruction algorithms. Database: MIT BIH normal sinus. (a) PRD performance for averaged over 18 ECG recordings of duration* 2 *hours [34]. (b) Box plots for all database records*

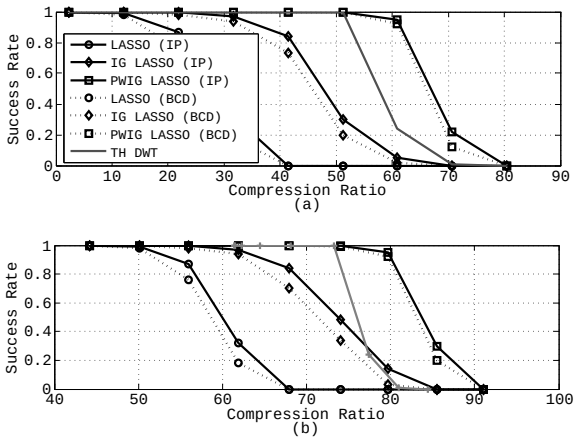



Figure 8: *Probability of recovering ECG segments with PRD less that 9%. Encoding method: sparse binary encoding with s:4. The CR has been evaluated at the output of (a) the random linear encoder, (b) the Huffman encoder. Database: MIT BIH normal sinus.*

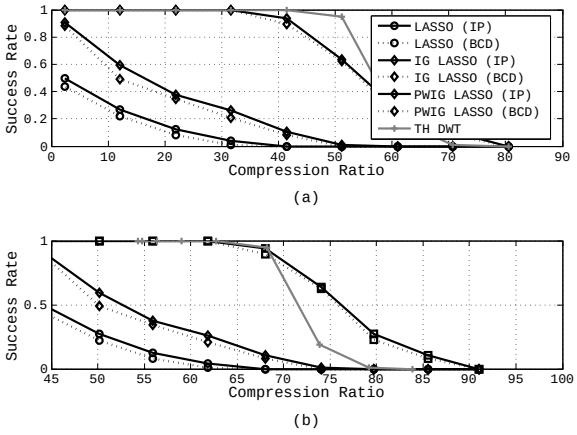



Figure 9: *Probability of recovering ECG segments with PRD ¡ 9%. Encoding method: sparse binary encoding with s:4. The CR has been evaluated at the output of (a) the random linear encoder (b) the huffman encoder. Database: MIT BIH compression.*

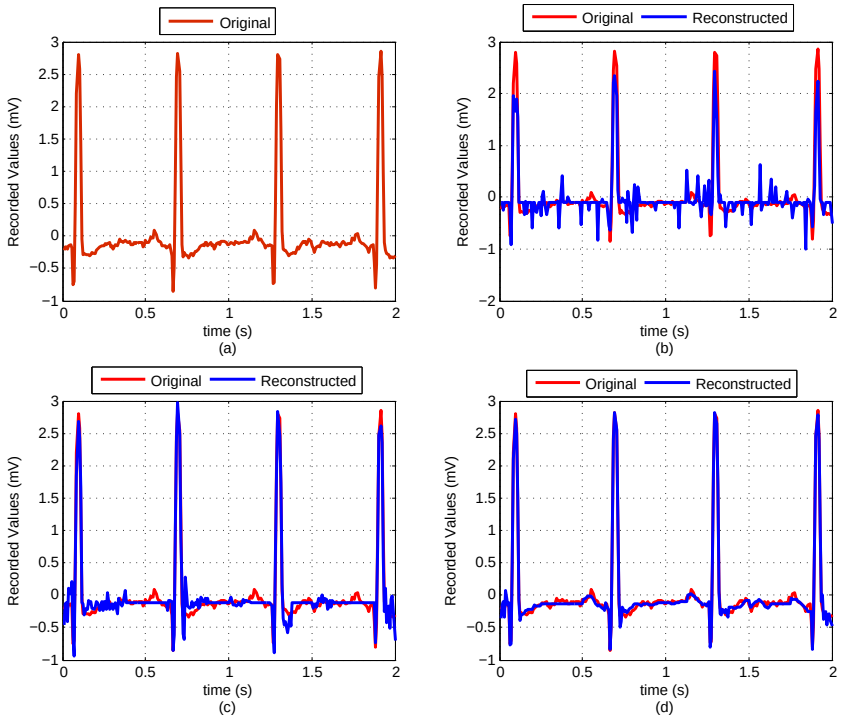



Figure 10: *Original and Reconstructed ECG data samples after applying estimators (ii), (iv) and (vi). Encoding method: sparse binary encoding with s:4. Number of encoded Data used for reconstruction* $K : 256$. *Database: MIT BIH normal sinus. (a) 2 sec of Original Signal, (b) Original vs Reconstructed from 56 % less data by LASSO (IP), (c) Original vs Reconstructed from 56 % less data by IG LASSO (IP), (d) Original vs Reconstructed from 56 % less data by PWIG LASSO (IP).*

is affected less than that of the IG LASSO. In addition, the best results are obtained for $d = 16$. In general, the choice of both the block length $d$ and the correlation factor $r$ are strongly connected to the sampling frequency that was used during the recording of the data. A rule of thumb that may be used for selecting those two parameters ($d$, $r$), is to check if, for these selected values the difference between the norms of non-zero blocks $\|\mathbf{x}[i]\|_2$ is minimized.

In fig. 7 (a) the obtained PRD, averaged over all the data of the MIT BIH normal sinus database, is plotted against the achieved compression ratio (CR) at the output of the random linear encoder, for the estimators: (i), (iii), (v) and (vii). The block length for the estimators (iii), (v) is selected to be equal to $d = 16$ and the correlation parameter for the estimator (v) is set to $r = 0.945$. The standard deviation of the PRD measurements is also provided in fig. 7 (b). On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, and the whiskers extend to the most extreme data points that are not considered outliers. By inspecting the plotted curves in figs. 7 (a), (b) we conclude again that CS algorithms that exploit the block sparsity of the ECG signal in the TD result in successful ECG signal recovery, by reducing the CR by 20% as compared to the achievable CR of the LASSO scheme. While the algorithms that further exploit the intra-block correlation can achieve up to 44% CR reduction.

Similar conclusions are drawn by inspecting also figs. 8 (a) and (b), where we plot the success rate versus the CR calculated either at the output of the random linear encoder (fig. 8 (a)) or at the output of the Huffman encoder (fig. 8 (b)). Note that the three additional encoding units that perform

inter-packet redundancy removal, quantization and Huffman coding (units with dashed lines in fig .2) at the transmitter increase the achieved CR by almost 20%. Moreover, the application of the BCD algorithm of section 4.3 for solving the SOCP problems result in similar performance with that achieved by the IP method.

Fig. 10 shows the original and the reconstructed segments that correspond to a 2 sec of ECG recordings taken from the Normal Sinus Rhythm Database. The number of samples used for recovery of the initial data was equal to $K = 112$, (which corresponds to 56% CR). Again it is shown that only estimators vi) result in a "Good Quality" ECG while the PRD of the estimators ii) and iv) is above 9%.

To consider more complex databases where type of arrythmias are also present we conducted experiments by using data taken from the MIT BIH compression database. In figs. 9 (a) and (b) we provide the results obtained by using all the data from the database, at the output of random linear encoder and Huffman encoder, respectively. It is clearly shown that the proposed scheme is even more robust in more complex ECG data bases as compared to the IG LASSO or LASSO approaches. In addition, as in the sinus rhythm case the application of the BCD as compared to the IP methods results in a similar performance at a lower reconstruction complexity. Finally, it should be noted that the proposed scheme especially in the case that the CR is computed at the output of the Huffman encoder, perform slightly better than the TH DWT method, but at a much lower compression cost, since the compression is achieved by simply generating random linear combinations of the recorded data.

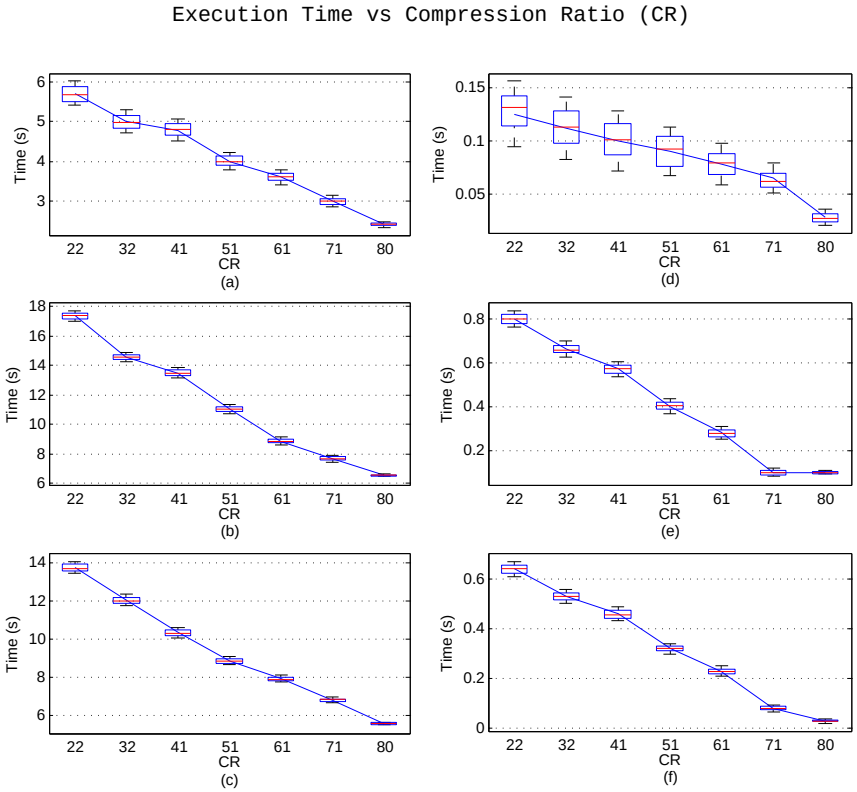


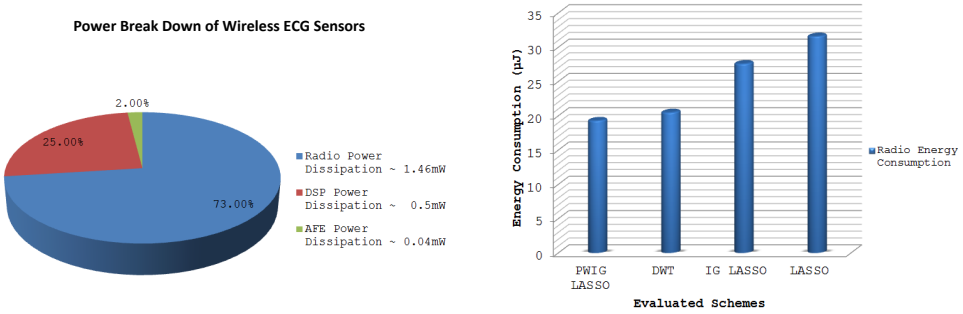

Figure 11: *Execution time of CS based estimators: (a) LASSO - IP (b) IG LASSO - IP (c) PWIG LASSO - IP (d) LASSO - BCD (e) IG LASSO - BCD (f) PWIG LASSO - BCD. Database: MIT BIH normal sinus.*

Fig. 11 gives the speed comparison of the 6 estimators (i) - (vi) that are executed on the receiver. The experiments were carried out on a computer with dual-core 2.9GHz CPU, 8 GB RAM and Windows 7 OS. Note that the execution time of PWIG LASSO is lower than that of IG LASSO (either when executing IP or when executing the BCD method), due to the fact that problem (16) requires less iterations to be executed (either IP or BCD iterations) as compared to that defined in eq. (10). The IG LASSO and PWIG LASSO schemes converge in only $K = 2 - 3$ iterations either when employing IP methods or the BCD algorithm of Section 4.3. The total number of iterations executed by the BCD in the PWIG LASSO and IG LASSO schemes was between $150 - 200$, while their execution time was 15-30 times lower than that of a standard IP method.

*5.3. Energy Consumption Gains*

In these subsection our goal is to provide some useful insights regarding the energy saving that can be achieved by the evaluated compression/ reconstruction schemes. In fig. 12 (a), we provide the power breakdown of existing low power ECG sensors according to [13], where it is shown that digital signal processing (DSP) and wireless data transmission dominates the system power consumption.

Based on [13], we assume that the transmit power of an ultra low power wireless ECG sensor is equal to 1.4 $mW$. Furthermore, fig. 8 (b), show that the PWIG LASSO method requires 75% less bytes, as compared to the original $N \times 12/8$ bytes, for ensuring an accurate reconstruction at the destination, while the DWT, the IG LASSO and the LASSO require 73%, 50%and 40% less bytes, respectively. To evaluate the energy consumed in the radio part




(a) (b)

Figure 12: (a) The power dissipation of the ECG wireless sensor nodes is dominated by DSP and wireless data transmission.(b) Radio Energy Consumption at the biosensor.

we select parameters that are in line with the IEEE 802.15.6 [41] standard specification[4], we assume that in each one of the aforementioned cases, the data required at the final destination, form a packet of length 108, 120, 204 and 248 bytes respectively, where the first 12 bytes correspond to the MAC and Physical Layer Preamble headers and the rest to the payload, consisting of the required encoded samples (Note that the payload data according to [41] vary between $0 - 255$ bytes). Assuming a data rate equal to 128 kbps, that is one of the supporting data rates in the 2360-2400 MHz band, the required times for transmitting each packet for each one of the evaluated methods is equal to 6.7 ms, 7.5 ms, 12.7 ms and 15.5 ms. Since we have assumed that processing is performed in blocks of samples that correspond to a signal observation of 1 s duration, for the rest time that the radio do not transmit, it is in a sleep mode consuming $P_s = 0.01mW$. In fig. 2 (b)

[4]IEEE 802.15.6 is a standard for short-range, wireless communications in the vicinity of, or inside, a human body

we present the energy consumption in the radio part of an ECG wireless sensor, for each one of the evaluated compression/reconstruction schemes. To evaluate the energy $E_R$, consumed in the radio part, we used the following equation $E_R = P_t T_p + P_s(1 - T_p)$ where $P_t$, $P_s$ denote the transmit and sleep power and $T_p$ is the time required for the transmission of a packet, consisting of the encoded ECG segment.

By inspecting figs. 12 (a), (b) it can be shown that the PWIG LASSO scheme, as compared to the IG LASSO and LASSO schemes, decreases the biosensor total energy consumption by 22% and 30% respectively. Recall that in all schemes the energy consumed in the DSP and AFE parts are the same, since the compression is performed in the same way. It should be also mentioned that although the radio energy consumption of the PWIG LASSO method is only 5% lower compared to the DWT method, the total energy gains are much higher, due to the fact that PWIG LASSO requires a simpler compression method to be executed at the transmitter. To be more specific, the authors in [R.16] showed that the LASSO method consume 30 times less energy in the DSP part of a Shimmer node, as compared to the DWT method. Based on that remark, the total energy savings of PWIG LASSO as compared to the DWT method due to the energy savings in the DSP part could be of the order of 26%, since the DSP part of a low power ECG sensor, consumes the 25% of the total energy consumption (fig. 12 (a)).

In a similar way, we can evaluate the energy consumed in the radio part of the receiver after taking into account that the receive power of an ultra low power wireless ECG sensor is equal to 1.6 $mW$. We expect that the PWIG LASSO scheme, as compared to the IG LASSO and LASSO schemes, will

decrease the receiver total energy consumption by factors larger than 22% and 30% respectively. This remark, is explained by the fact that apart from the gains offered by the achieved CR in the radio part of the receiver, the exploitation of block sparsity offers significant gains at the DSP part of the receiver, according to the results presented in fig. 11.

## 6. Concluding Remarks

ECG telemonitoring via WBANs introduces several challenges, that motivate the design of ECG compression schemes, with reduced compression and reconstruction computational requirements at both the transmitter and receiver. The application of the theory of Compressed Sensing (CS) to the problem at hand, can achieve those requirements. We showed via extensive simulations that the exploitation of the block structure of ECG signal in the TD, and more importantly in a specific UD can lead to significant savings in the amount of data that should be transmitted from the biosensor in order to achieve an accurate signal reconstruction at any receiver in the WBAN (i.e., smart phone or remote terminal). The proposed algorithms for ECG signal recovery at the destination, achieve improved reconstruction capabilities with significant gains in both compression ratio and computational cost. Although, the focus was the wireless ECG telemonitoring, the proposed algorithms are applicable to other telemedicine applications such as telemonitoring of electroencephalograph [39], and electromyography [17] where the signals have similar block sparse structure either in the TD or in other transform domains.

## 7. References


[1] H. Cao, V. Leung, C. Chow, and H. Chan, "Enabling technologies for wireless body area networks: A survey and outlook," *IEEE Communications Magazine*, vol. 47, no. 12, pp. 84–93, 2009.

[2] S. S. Mahmoud, Q. Fang, Z. M. Hussain, I. Cosic, "A blind equalization algorithm for biological signals transmission," *Digital Signal Processing*, vol. 22, no. 1, pp. 114–123, 2012.

[3] J. Pandey and B. Otis, "A sub-100 $\mu$ w mics/ism band transmitter based on injection-locking and frequency multiplication," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 5, pp. 1049–1058, 2011.

[4] S. M. S. Jalaleddine, C. Hutchens, R. Strattan, and W. Coberly, "Ecg data compression techniques-a unified approach," *IEEE Transactions on Biomedical Engineering*, vol. 37, no. 4, pp. 329–343, 1990.

[5] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge Univ. Press, 2004.

[6] L. Sornmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. Elsevier/Academic press: New York, 2005.

[7] M. Abo-Zahhad, A. F. Al-Ajlouni, S. M. Ahmed, R. J. Schilling, "A new algorithm for the compression of ECG signals based on mother wavelet parameterization and best-threshold levels selection," *Digital Signal Processing*, vol. 23, no. 3, pp. 1002–1011, 2013.

[8] M. Hilton, "Wavelet and wavelet packets compression of electrocardiogram," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 5, pp. 394–402, 1997.

[9] Z. Lu, D. Y. Kim, and W. A. Pearlman, "Wavelet compression of ECG signals by the set partitioning in hierarchical trees algorithm," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 7, pp. 849–856, 2000.

[10] R. Benzid, F. Marir, A. Boussaad, M. Benyoucef, and D. Arar, "Fixed percentage of wavelet coefficients to be zeroed for ECG compression," *Electronic Letters*, vol. 39, no. 11, pp. 830–831, 2003.

[11] Y. Eldar, "Compressed sensing of analog signals in shift-invariant spaces," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 2986–2997, 2009.

[12] R. Agarwal and S. Sonkusale, "Input-feature correlated asynchronous analog to information converter for ecg monitoring," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 5, no. 5, pp. 459–467, 2011.

[13] R. Yazicioglu, S. Kim, T. Torfs, H. Kim, and C. Van Hoof, "A 30 $\mu$ w analog signal processor asic for portable biopotential signal monitoring," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, pp. 209–223, 2011.

[14] M. Mishali, Y. Eldar, O. Dounaevsky, and E. Shoshan, "Xampling: Analog to digital at sub-nyquist rates," *IET Circuits, Devices Systems*, vol. 5, no. 1, pp. 8–20, 2011.

[15] E. C. Pinheiro, O. A. Postolache, and P. S. Girão, "Implementation of compressed sensing in telecardiology sensor networks," *International Journal of Telemedicine Applications*, vol. 2010, pp. 71–77, 2010.

[16] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vandergheynst, "Compressed sensing for real-time energy-efficient ecg compression on wireless body sensor nodes," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 9, pp. 2456–2466, 2011.

[17] A. Dixon, E. Allstot, D. Gangopadhyay, and D. Allstot, "Compressed sensing system considerations for ecg and emg wireless biosensors," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 6, no. 2, pp. 156–166, 2012.

[18] Z. Zhang, T.-P. Jung, S. Makeig, and B. Rao, "Compressed sensing for energy-efficient wireless telemonitoring of noninvasive fetal ecg via block sparse bayesian learning," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 2, pp. 300–309, 2013.

[19] F. Chen, A. Chandrakasan, and V. Stojanovic, "Design and analysis of a hardware-efficient compressed sensing architecture for data compression in wireless sensors," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 3, pp. 744–756, 2012.

[20] G. Friesen, T. Jannett, M. Jadallah, S. Yates, S. Quint, and H. Nagle, "A comparison of the noise sensitivity of nine qrs detection algorithms," *IEEE Transactions on Biomedical Engineering*, vol. 37, no. 1, pp. 85–98, 1990.

[21] J. Friedman, T. Hastie, H. Hofling, and R. Tibshirani, "Pathwise coordinate optimization," Annals of Applied Statistics, Tech. Rep., 2007.

[22] A. Rakotomamonjy, "Surveying and comparing simultaneous sparse approximation (or group-lasso) algorithms", *Signal Processing*, vol. 91, no. 7, pp. 1505–1526, 2011.

[23] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

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

[25] E. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[26] M. Stojnic and F. Parvaresh and B. Hassibi, "On the Reconstruction of Block-Sparse Signals With an Optimal Number of Measurements," *IEEE Transactions on Signal Processing*, vol.57, no.8, pp.3075–3085, 2009.

[27] R. Khandpur, *Handbook of Biomedical Instrumentation.* Tata McGraw-Hill Education, 2003.

[28] E. Candes and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[29] A. Gilbert, P. Indyk, "Sparse Recovery Using Sparse Matrices," *Proceedings of the IEEE*, vol.98, no.6, pp.937–947, 2010.

[30] F. Alizadeh and D. Goldfarb, *Second-order cone programming.* Math. Program., 2003, no. 95.

[31] E. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted $\ell_1$ minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5-6, pp. 877–905, 2008.

[32] K. Lange, D. R. Hunter, and I. Yang, "Optimization Transfer Using Surrogate Objective Functions," *Journal of Computational and Graphical Statistics*, vol. 9, pp. 1–20, 2000.

[33] D. Wipf and S. Nagarajan, "Iterative reweighted $\ell_1$ and $\ell_2$ methods for finding sparse solutions," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 317–329, 2010.

[34] A. L. Goldberger and L. A. N. Amaral and L. Glass and J. M. Hausdorff and P. Ch. Ivanov and R. G. Mark and J. E. Mietus and G. B. Moody and C. Peng and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 100, no. 23, pp. e215–e220, 2013.

[35] M. Lobo and L. Vandenberghe and S. Boyd and H. Lebret, "Applications of second-order cone programming;," *Linear Algebra Applications*, vol. 284, pp. 193–228, 1998.

[36] Y. Zigel, A. Cohen, and A. Katz, "The weighted diagnostic distortion (wdd) measure for ecg signal compression," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 11, pp. 1422–1430, 2000.

[37] D. Angelosante, J. Bazerque, and G. Giannakis, "Online adaptive estimation of sparse signals: Where rls meets the $\ell_1$ -norm," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3436–3447, 2010.

[38] H. Zou, "The Adaptive Lasso and Its Oracle Properties," *Journal of the American Statistical Association*, vol. 101, no. 476, pp. 1418–1429, 2006.

[39] Z. Zhang, T.-P. Jung, S. Makeig, and B. Rao, "Compressed sensing of eeg for wireless telemonitoring with low energy consumption and inexpensive hardware," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 1, pp. 221–224, 2013.

[40] A.H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley, New York, 2003.

[41] LAN/MAN Standards Committee of the IEEE Computer Society, "IEEE Standard for Local and metropolitan area networks - Part 15.6: Wireless Body Area Networks," *IEEE Std 802.15.6-2012*, Feb. 2012.

**Aris S. Lalos** received the Ph.D. degree in signal processing for wireless communications from the Computer Engineering and Informatics Department (CEID), School of Engineering (SE), University of Patras (UoP), Rio-Patras, Greece, in 2010. He has been a research fellow at Signal Processing and Communications Laboratory, CEID, SE, UoP, Rio-Patras, Greece

from 2005 to 2010. In period 2010-2011 was a telecommunication research engineer at Analogies S.A, an early stage start up. He is currently a postdoctoral researcher at signal theory and communication department, Technical University of Catalonia, Barcelona, Spain.

**Luis Alonso** received the PhD from UPC (Barcelona) in 2001 and got a permanent position at UPC becoming an Associate Professor in 2006. Co-founder of the Wireless Communications and Technologies Research Group (WiComTec), to which currently belongs. His current research interests are within the field of medium access protocols, radio resource management, cross-layer optimization, cooperative transmissions, cognitive radio and QoS features for all kind of wireless communications systems. He is author of forty research papers, one book, twelve chapters of books and more than one hundred papers in international congresses and symposiums and received several best paper awards.

**Christos Verikoukis** received his Ph.D. from the Technical University of Catalonia in 2000. He is currently the Head of the SMARTECH department at CTTC and he is an adjunct associate professor in the area of ICT for Healthcare at Barcelona University (UB). He has published 55 journal papers and over 120 conference papers. He is also the co-author of 2 books, 14 chapters of different books and of 2 patents. Dr. Verikoukis has participated in more than 30 competitive research projects. He is currently an officer (the secretary) of the IEEE ComSoc Technical Committee on Communication Systems Integration and Modeling (CSIM).

Table 1: Summary of the Iterative Reweighed Group LASSO (IG LASSO) and Prewhitened Iterative Reweighed Group LASSO (PWIG LASSO) Algorithms

**(PW)IG LASSO:**

**Inputs:** Encoding Matrix: $\mathbf{A}$

Received packets: $\mathbf{y}$, Block length $d$

**Output:** Estimated ECG Data $\mathbf{x}$, $dc$ coefficient

**1. Initialize $\mathbf{y}_n$, $\mathbf{A}_n$, $w_i^{(0)}$**

$\mathbf{y}_n = \left(I_M - \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}\right)\mathbf{y}$,$\mathbf{A}_n = \left(I_M - \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}\right)\mathbf{A}$, $\mathbf{a} = \mathbf{A}\mathbf{1}_N$

Solve $\mathbf{x}^{(0)} := \arg\min_x \|\mathbf{y}_n - \mathbf{A}_n\mathbf{x}\|_2^2 + \sum_{i=1}^{R} \lambda\|\mathbf{x}[i]\|_2$

by i) interior point methods (IP) or ii) Algorithm BCD of Section 4.3.

$w_i^{(0)} := \left(\|\mathbf{x}^0[i]\|_2 + \epsilon\right)^{-1}, \quad i = 1, \ldots, R$

**2. Evaluate $\mathbf{T}_i$, $\forall i$ from eqs. (13)***

**3. Initialize $\mathbf{s}^{(0)}$ and compute matrix $\mathbf{\Psi}$**

$\mathbf{s}[i] = \mathbf{T}_i^{-1/2}\mathbf{x}[i]$, $\mathbf{s} = \left[\mathbf{s}^T[1]\ \ \mathbf{s}^T[2]\ \ \ldots \mathbf{s}^T[R]\right]$

$\mathbf{\Psi} = \mathbf{A_n} diag\left\{\mathbf{T}_1^{1/2}, \ldots, \mathbf{T}_R^{1/2}\right\}$

**4. repeat for each iteration $l = 1, \ldots, K$**

a. Solve problem (16) by either applying

i) interior point methods (IP) or ii) Algorithm BCD of Section 4.3.

b. Evaluate $\mathbf{x}^{(l)}[i] = \mathbf{T}_i^{1/2}\mathbf{s}^{(l)}[i] \forall\ i$

c. Update weights $w_i^{(l+1)}$ according to eq. (17)

d. Update $\mathbf{T}_i$, $\forall i$ from eqs. (13)**

e. Update $\mathbf{\Psi} = \mathbf{A_n} diag\left\{\mathbf{T}_1^{1/2}, \ldots, \mathbf{T}_R^{1/2}\right\}$**

**end repeat**

**5. Estimate dc offset coefficient**

$\hat{d}_c = \left(\mathbf{a}^T\mathbf{a}\right)^{-1}\mathbf{a}^T\left(\mathbf{y} - \mathbf{A}\mathbf{x}^{(K)}\right)$

** By setting $T_i = \mathbf{I}_d$, $\forall i = 1, \ldots, R$ the PWIG LASSO scheme reduces to the IG LASSO scheme that exploits block sparsity in the TD.*

*** Steps d,e can be eliminated by using a fixed $r$ at each iteration.*