IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

A Fast Compressed Sensing Decoding Technique for Remote ECG Monitoring Systems

AHMAD M. RATEB, (Member, IEEE)
Faculty of Electronic and Information Engineering, Huaiyin Institute of Technology, Huaian, Jiangsu 223002 China.
e-mail: ahmadmrateb@ieee.org

ABSTRACT Compressed Sensing (CS) has been proposed as a low-complexity ECG data compression scheme for wearable wireless bio-sensor devices. However, CS decoding is characterized by high computational complexity. As a result, it represents a burden to the computational and energy resources of the network gateway node, where decoding is performed. In this paper, we propose a Fast Compressive Electrocardiography (FCE) technique to address this problem. CS decoding in FCE is based on Weighted Regularized Least-Squares (WRLS), rather than the standard approach based on ℓ_1 norm minimization. The WRLS formulation takes into account prior knowledge of ECG signal properties to estimate an optimally compact and accurate representation of ECG signals. Numerical results show that decoding by FCE is on average 33 times faster than the fastest tested CS-based ECG decoding technique. In addition, high-quality ECG signal reconstruction by FCE is achieved at 32% higher compression ratio. Therefore, FCE can contribute to improving the overall energy and computational resource efficiency of CS-based remote ECG monitoring systems.

INDEX TERMS compressed sensing (CS), electrocardiogram (ECG), random demodulator, remote health monitoring, wireless body-sensor network (WBSN).

I. INTRODUCTION

Remote health monitoring systems have recently gained significant importance, due to their role in treatment, prevention and early detection of diseases. Specifically, remote monitoring of the heart Electrocardiogram (ECG) is of prime interest, since it requires continuous and long-term monitoring [1], [2]. Fig. 1 illustrates an overview of an exemplary remote ECG monitoring system. In the context of a Wireless Body-Sensor Network (WBSN), a wearable ECG *sensor node* acquires, compresses, and wirelessly transmits the patient's ECG data. A *gateway node* receives, reconstructs and processes the ECG signals to extract useful medical data, such as heart rate, rhythm and various indicative intervals [3], [4]. It then forwards these data to a cloud-based database that can be accessed by medical specialists. Recent works have shown that ECG data processing and information extraction at the gateway node is more energy-efficient than blindly forwarding raw ECG data to the cloud. In addition, it reduces the traffic load on the network [1], [5], [6]. The gateway node maintains the option of forwarding selected abnormal ECG segments to be examined by the medical specialist.

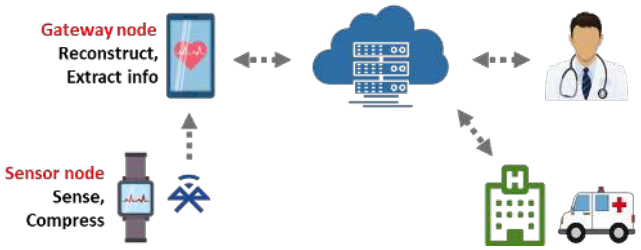


FIGURE 1: An overview of an exemplary remote ECG monitoring system.

Due to the restrictions on its size, weight and cost, the sensor node has a limited battery capacity. The majority of the energy consumed at the sensor node is referred to wireless transmission, followed by digital signal processing [7], [8]. Consequently, Compressed Sensing (CS) [9] has been proposed as a low-complexity ECG data compression scheme to improve the energy efficiency of the sensor node [10]. CS encoding is simply achieved by linearly projecting each frame of acquired ECG samples to a random *sensing matrix*, yielding a smaller sized compressed frame. Standard

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3035423, IEEE Access

CS decoding is based on solving an ℓ_1 norm minimization problem via convex optimization methods, or via greedy algorithm implementations [11], [12]. The seminal work in [13] compared CS-based ECG (CS-ECG) compression to standard ECG compression based on Discrete Wavelet Transform (DWT). The results confirmed the effectiveness of CS-ECG in terms of energy efficiency, in return for reduced reconstructed signal quality and higher decoding complexity. Similar results were reported in [14]–[16].

Most of the recent works on CS-ECG systems focused on improving the reconstructed signal quality through adapting the CS encoder to the sensed ECG signal. This was mainly achieved by optimizing the structure of the sensing matrix to benefit from structural, temporal and statistical properties of ECG signals. These works are well reviewed in the recent survey of adapted CS techniques in [17]. However, only a few works focused on performing this adaptation at the CS decoder side. In [18], the block-sparse structure and inter-block correlations of ECG signals were exploited for improving the reconstructed signal quality by using the Block-Sparse Bayesian Learning (BSBL) algorithm [19].

In [20], a class of artificial neural networks called Restricted Boltzmann Machines (RBM) were employed to create a statistical model of the ECG signal sparsity pattern. This model is plugged into the CS decoder to improve compression ratio and reconstruction performance. However, this approach has a large impact on increasing decoding complexity. A Weighted ℓ_1 Minimization (WLM) technique was proposed in [21]. Prior knowledge of the wavelet coefficients decay factors of ECG signals was incorporated into the ℓ_1 minimization problem used for CS decoding. This enabled the decoder to identify significant DWT coefficients more accurately. Orthogonal Matching Pursuit (OMP) [22] is a widely used fast CS decoding algorithm. Weighted OMP (WOMP) in [23] adapted the WLM technique to OMP in order to benefit from the speed advantage of OMP.

The aforementioned adapted CS-ECG decoding techniques achieved considerable gains in terms of minimizing the reconstructed ECG signal error. However, the computational complexity of these techniques is at least as high as standard CS decoding. In most cases, the patient's smartphone plays the role of the WBSN gateway node, since most patients are not expected to acquire specialized devices. Thus, CS-ECG decoding at the gateway node could represent a considerable computational burden. This leads to fast battery drainage, especially for the cases where long-term monitoring is necessary. In addition, it may not be possible reconstruct the signals in the real time, especially if large ECG frame length was used [5], [6], [24].

In this paper, we propose a *Fast Compressive Electrocardiography* (FCE) technique, which mainly addresses the problem of high decoding complexity of the state-of-the-art CS-ECG systems. In addition, it contributes to improving the reconstructed ECG signal quality. FCE is an adapted CS decoding technique based on Weighted Regularized Least Squares (WRLS). WRLS is known to be significantly less complex than ℓ_1 norm minimization [25]. In FCE, the WRLS problem targets estimating a small set of Discrete Cosine Transform (DCT) coefficients, which accurately represent the sensed ECG signal. Prior knowledge of the decay profile of these coefficients is exploited to assign their respective weights. In addition, the values of tunable parameters in the WRLS problem are optimized to maximize the reconstructed signal quality. Nevertheless, FCE does not require any modification to the CS encoder.

CS decoding by FCE is achieved with significantly lower computational complexity than all current CS-ECG decoding techniques. This will be shown both analytically and experimentally throughout this paper. As a result, the impact of the ECG monitoring application on the patient's smartphone battery life can be minimized. In addition, real-time ECG signal reconstruction can be practically feasible, even for large signal frame sizes. Furthermore, we will show that FCE achieves high-quality reconstruction at relatively higher compression ratios, and hence more energy savings are achievable at the sensor node.

The remaining part of this paper is organized as follows: in Section II, we provide an essential background on CS. We introduce the FCE technique in Section III. We then present our numerical experiments and discuss their results in Section IV. Finally, we draw our conclusions in Section V.

II. BACKGROUND

A. CS FRAMEWORK

In a digital CS paradigm as adopted in [13], the analog ECG signal picked up by the ECG sensor is input to an ADC operating at f_s samples/second. The ADC outputs a frame of N samples every sensing interval $T = N/f_s$ seconds. Suppose the ECG frame is represented by a column vector $\mathbf{x} \in \mathbb{R}^N$. CS encoding is a simple linear transformation, which can be modeled by the following matrix equation:

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \mathbf{q} \tag{1}$$

where $\mathbf{\Phi} \in \mathbb{R}^{M\times N}$ is the sensing matrix, $\mathbf{y} \in \mathbb{R}^M$ is called the *measurements vector*, and $\mathbf{q} \in \mathbb{R}^M$ represents a bounded error term ($\|\mathbf{q}\|_2 \leq \varepsilon$) that may be present due to quantization and additive noise. Compression performance is quantified by the Compression Ratio (CR), which is defined as:

$$\mathrm{CR} \triangleq \frac{N-M}{N} \times 100\% \tag{2}$$

From the basic theory of signal analysis [26], any signal $\mathbf{x} \in \mathbb{R}^N$ can expressed as a weighted sum of the columns of an orthonormal (unitary) *basis matrix* $\mathbf{\Psi} \in \mathbb{R}^{N\times N}$ as follows:

$$\mathbf{x} = \mathbf{\Psi}\mathbf{u} \tag{3}$$

where $\mathbf{u} \in \mathbb{R}^N$ is referred to as the *representation* of $\mathbf{x}$ in the domain of the basis $\mathbf{\Psi}$. A signal $\mathbf{x}$ is called $s-$*sparse* in the domain of $\mathbf{\Psi}$ if the number of non-zero coefficients (or those of considerable magnitude) in $\mathbf{u}$ is at most s, such that $s \ll N$. DWT bases have been reported to provide

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3035423, IEEE Access

highly sparse representations of ECG signals compared to other bases. Specifically, Daubechies wavelets (Db-4 and Db-10) [13], [27], Symlets (Sym-6) [21] and Coiflets (Coif-5) [28].

By combining (1) and (3), we can directly link the compressed measurements vector $\mathbf{y}$ to the sparse vector $\mathbf{u}$, as follows:

$$\mathbf{y} = \mathbf{\Phi}\mathbf{\Psi}\mathbf{u} + \mathbf{q} = \mathbf{H}\mathbf{u} + \mathbf{q} \tag{4}$$

where $\mathbf{H} = \mathbf{\Phi}\mathbf{\Psi} \in \mathbb{R}^{M\times N}$ is called the *compression matrix*. The reconstructed sparse vector $\hat{\mathbf{u}}$ can be found by solving the following convex optimization problem, which is known as *Basis Pursuit Denoising* (BPDN) [29]:

$$\hat{\mathbf{u}} = \arg\min_{\mathbf{u}} \|\mathbf{u}\|_1 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{H}\mathbf{u}\|_2 \leq \varepsilon \tag{5}$$

where $\|\mathbf{u}\|_1 = \sum_{i=0}^{N-1} |u_i|$ is the ℓ_1 norm of the vector $\mathbf{u}$. Minimizing the ℓ_1 norm serves as a proxy for promoting the sparsity of the solution [29]. A plenty of methods can be used for solving the BPDN problem efficiently, such as interior-point methods [30] and Spectral Projected-Gradient (SPG) method [31]. In addition, the solution can be approximated by fast greedy iterative algorithms, such as OMP [22] and Compressed Sampling Matching Pursuit (CoSaMP) [32]. CS decoding algorithms are well reviewed in [11], [12].

Since (4) is an under-determined system of linear equations, it admits an infinite number of feasible solutions. In order for (5) to yield an accurate estimate of the sparse unknown vector, the compression matrix $\mathbf{H}$ should satisfy a *Restricted Isometry Property* (RIP) [33]. The RIP is satisfied with very high probability for sensing matrices with iid random entries of Gaussian or Bernoulli (± 1) distributions. This is valid over any orthonormal basis, provided that $M \geq c_0 s \log(N/s)$, where c_0 is a numerical constant [34]. Acquiring insufficient number of measurements leads to considerable reconstruction error [35], [36].

The RIP provides strong recovery guarantees. However, verifying the RIP for arbitrary compression matrices is highly complicated. Alternatively, more computationally tractable guarantees are provided by the *mutual coherence* of the compression matrix, given by [37]:

$$\mu(\mathbf{H}) = \max_{1\leq i\neq j\leq N} \frac{|\mathbf{h}_i^T \mathbf{h}_j|}{\|\mathbf{h}_i\|_2 \|\mathbf{h}_j\|_2} \tag{6}$$

where $\mathbf{h}_i$ and $\mathbf{h}_j$ are the i-th and the j-th columns of $\mathbf{H}$ respectively. The equation $\mathbf{y} = \mathbf{H}\mathbf{u}$ has a unique s-sparse solution if $\mu < 1/(2s-1)$, where $\sqrt{(N-M)/(N-1)M} \leq \mu \leq 1$ [38]. Extended guarantees for the noisy case in (4) were provided in [39]. In general, a smaller value of μ corresponds to higher probability of accurate reconstruction of $\mathbf{u}$ by solving (5).

B. LOW-COMPLEXITY CS ENCODING

CS encoding is achieved in the digital domain by applying the sensing matrix to the data vector, which requires $\mathcal{O}(MN)$ operations. For large N, such process could be computationally demanding for the sensor node. In effort to reduce the encoding complexity, the Random Sparse Binary Matrix (RSBM) was proposed in [13], such that each column of the sensing matrix contains exactly $d \ll N$ randomly located 1's. It was shown that the mutual coherence of RSBMs with DWT (db10) basis can be as low as the mutual coherence of the Gaussian random sensing matrix (which is usually used as a reference) for $d \geq 12$ [13].

One of the fundamental strengths of CS is that its encoding stage can be performed entirely in the analog domain. The measurements vector can be directly acquired by an ADC operating at a low sampling rate (which could go below Nyquist rate). The Random Demodulator (RD) shown in Fig. 2 is one of the simplest and most energy-efficient analog CS acquisition architectures [40]–[42]. The input analog signal $x(t)$ is multiplied by a continuous-time pseudo-random chipping sequence $p(t) = \pm 1$ by using an analog mixer. The chipping rate of $p(t)$ is set to the original signal sampling rate: f_s chips per second. Assuming that N is divisible by M, an integrator accumulates the product signal, and the output is sampled every $L = N/M > 1$ chips, with resetting the integrator. Thus, the ADC sampling rate is reduced to f_s/L samples per second.

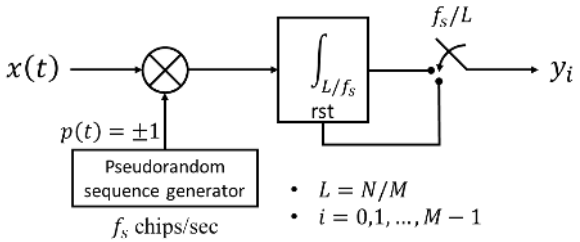


FIGURE 2: Block diagram of the random demodulator (RD) architecture [42].

CS encoding by the RD can also be modeled by (1) [42]. Suppose every N chips of $p(t)$ are divided into M consecutive segments. Each segment is a row-vector of L antipodal entries. The i-th segment is given by: $\mathbf{p}_i = \{p_{(i,0)}, p_{(i,1)}, \ldots, p_{(i,L-1)}\}$, where $i = 0, 1, \ldots, M-1$. The RD sensing matrix is hence modeled as a block-diagonal matrix of these segments, which is expressed as [42]:

$$\mathbf{\Phi} = f_s^{-1} \cdot \text{blkdiag}(\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_{M-1}) \tag{7}$$

The RD was originally proposed for band-limited multi-tone signals, which are sparse over the DFT basis. However, it was found to be highly coherent with DWT bases (i.e. μ is large), which made it a bad choice for CS-ECG systems [13].

In effort to address this problem, other analog CS-ECG encoding architectures were proposed such as the Random Demodulator Pre-Integrator (RMPI) [43], [44], Spread-spectrum Random Demodulator Pre-Integrator (SRMPI) [45] and Compressed Sensing Analog Front-End (CS-AFE) [46]. All these architectures are more or less based on using several RD channels in parallel. Thus, the implementation size, energy consumption and cost incurred with such architectures is at least several times higher than the RD.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3035423, IEEE Access

III. PROPOSED TECHNIQUE

A. MOTIVATION

Consider an ECG signal frame of N samples selected arbitrarily from the MIT-BIH arrhythmia database [47], [48]. Fig. 3 compares the magnitudes of its DWT representation coefficients to their DCT counterparts. We observe that although the DCT representation is less sparse than the DWT representation (i.e. has more non-zeros), significant DCT coefficients are concentrated within the lower 30% of indices. This is referred to the *energy compaction* property of DCT [26], [49]. Based on this property, the DCT representation of a finite sequence is often more concentrated at low indices compared to other transforms. In addition, ECG signals can be well approximated by the first $k < N$ coefficients of the DCT representation, where approximation error rapidly decays with k. This is referred to the exponential coefficient magnitude decay profile, which can be observed in Fig. 3.

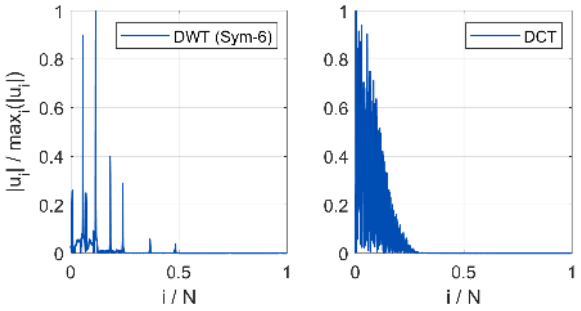


FIGURE 3: Normalized coefficient magnitudes versus coefficient index for DWT (Sym-6 wavelet with 3 levels of decomposition) and DCT basis representations. ECG frame length is $N = 1024$ taken from record no. 123 of the MIT-BIH arrhythmia database.

Based on the above discussion, it is possible to reconstruct the ECG signal with arbitrarily small error by targeting estimating a sufficient number of low-index DCT coefficients. In addition, since the coefficients within the low-index region are non-sparse, computationally demanding ℓ_1 minimization approach is no more necessary. It can be replaced by simpler and more flexible ℓ_2 norm minimization. However, the success of this approach is bound to proper exploitation of the prior knowledge of the coefficient decay profile, as to be shown next.

B. METHOD

Let the basis matrix $\mathbf{\Psi}$ be set as the orthonormal inverse DCT-2 matrix, whose elements are given by [26], [49]:

$$\psi_{i,j} = \frac{\beta(j)}{\sqrt{N}} \cos\left(\frac{(2j+1)i\pi}{2N}\right), \quad i,j = 0,1,\ldots,N-1 \tag{8}$$

where $\beta(j) = 1$ for $j = 0$, and equals $\sqrt{2}$ otherwise. Thus, for a given value of $k < N$, the ECG signal can be expressed as:

$$\mathbf{x} = \begin{bmatrix}\mathbf{\Psi}_k & \tilde{\mathbf{\Psi}}_k\end{bmatrix}\begin{bmatrix}\mathbf{u}_k \\ \tilde{\mathbf{u}}_k\end{bmatrix} = \mathbf{\Psi}_k\mathbf{u}_k + \tilde{\mathbf{\Psi}}_k\tilde{\mathbf{u}}_k \tag{9}$$

where $\mathbf{\Psi}_k \in \mathbb{R}^{N\times k}$ is the collection of the first k columns of $\mathbf{\Psi}$, and $\tilde{\mathbf{\Psi}}_k \in \mathbb{R}^{N\times(N-k)}$ is the collection of the remaining columns of the matrix. Similarly, $\mathbf{u}_k \in \mathbb{R}^k$ consists of the first k entries of $\mathbf{u}$, while $\tilde{\mathbf{u}}_k \in \mathbb{R}^{N-k}$ consists of the remaining entries.

Based on (9), the measurements vector can be written as $\mathbf{y} = \mathbf{\Phi}(\mathbf{\Psi}_k\mathbf{u}_k + \tilde{\mathbf{\Psi}}_k\tilde{\mathbf{u}}_k) + \mathbf{q}$. Suppose we choose k such that $\|\mathbf{\Phi}\tilde{\mathbf{\Psi}}_k\tilde{\mathbf{u}}_k\|_2 \ll \|\mathbf{\Phi}\mathbf{\Psi}_k\mathbf{u}_k\|_2$. Hence, we may treat the term $\mathbf{v} = \mathbf{\Phi}\tilde{\mathbf{\Psi}}_k\tilde{\mathbf{u}}_k \in \mathbb{R}^{M\times 1}$ as a small additive error/noise vector as follows:

$$\mathbf{y} = \mathbf{\Phi}\mathbf{\Psi}_k\mathbf{u}_k + \mathbf{v} + \mathbf{q} = \mathbf{H}_k\mathbf{u}_k + (\mathbf{v} + \mathbf{q}) \tag{10}$$

where $\mathbf{H}_k = \mathbf{\Phi}\mathbf{\Psi}_k \in \mathbb{R}^{M\times k}$ consists of the first k columns of $\mathbf{H}$. By targeting estimating the non-sparse vector $\mathbf{u}_k$, we have eliminated the need for estimating the support (non-zeros locations) of the entire vector $\mathbf{u}$. Thus, $\mathbf{u}_k$ can be simply estimated by posing the following Least Squares (LS) problem:

$$\hat{\mathbf{u}}_k = \arg\min_{\mathbf{u}_k} \|\mathbf{y} - \mathbf{H}_k\mathbf{u}_k\|_2^2 \tag{11}$$

The solution of above problem is given by [25]:

$$\hat{\mathbf{u}}_k = (\mathbf{H}_k^T\mathbf{H}_k)^{-1}\mathbf{H}_k^T\mathbf{y}, \quad k \le M \tag{12}$$

The solution in (12) is only valid when the system of equations in (10) is fully determined or over-determined, i.e. for $k \le M$. There is no closed-form expression of the solution for the under-determined case: $k > M$. However, for high compression ratios (i.e. $M \ll N$), there is a high probability that $k > M$. Thus, it is of high importance to find an optimal solution for all values of k, especially for $k > M$. To alleviate this problem, we employ Weighted Regularized Least-Squares (WRLS) [25], which adds a second objective to (11). The trade-off between the two objectives is controlled by the regularization parameter $\lambda > 0$, as follows:

$$\hat{\mathbf{u}}_k = \arg\min_{\mathbf{u}_k} \|\mathbf{y} - \mathbf{H}_k\mathbf{u}_k\|_2^2 + \lambda\|\mathbf{W}_k\mathbf{u}_k\|_2^2 \tag{13}$$

In the above formulation, we target minimizing the energy of a weighted version of $\mathbf{u}_k$ as a second objective, where $\mathbf{W}_k = \mathrm{diag}(w_0, w_1, \ldots, w_{k-1})$ is a $k\times k$ diagonal weighting matrix. The weighting matrix will enable us to incorporate the coefficients decay information into the WRLS problem, in order to further improve decoding performance. The problem in (13) has a closed-form solution given by [25]:

$$\hat{\mathbf{u}}_k = (\mathbf{H}_k^T\mathbf{H}_k + \lambda\mathbf{W}_k^2)^{-1}\mathbf{H}_k^T\mathbf{y}, \quad 1 \le k \le N \tag{14}$$

This solution is valid for all feasible values of k and M, due to the presence of the diagonal matrix $\mathbf{W}_k^2$ within the matrix inversion.

Proper design of the weighting matrix require a mathematical model of the coefficients magnitude decay profile. The root-mean-square (RMS) values of the DCT coefficients of ECG signals are plotted versus their index in Fig. 4. These values were computed over the collection of full-length

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3035423, IEEE Access

records no. 100, 112, 123, 200 and 222 of the MIT-BIH database. We modeled these data by the following function:

$$f(i) = \exp\left[-a_1 \sin(b_1 q(i) + c_1) - a_2 \sin(b_2 q(i) + c_2)\right] \tag{15}$$

where $q(i) = (i+1)/N$ for all $i = 0, 1, \ldots, N-1$. The coefficients are given by: $a_1 = 13.7$, $b_1 = 1.35$, $c_1 = 0.06$, $a_2 = 0.65$, $b_2 = 20.45$ and $c_2 = 1.42$. This model was produced by fitting the sum of two sine functions to the log of the data points, with higher weights assigned to low-index coefficients due to their significance. The curve produced by the above model is also shown in Fig. 4.

Minimizing the energy of the term $\mathbf{W}_k \mathbf{u}_k$ in (13) implies that the weighting coefficients should equalize the decaying magnitude profile of $\mathbf{u}_k$. This can be achieved by setting the i-th entry of the weighting matrix diagonal w_i as the normalized reciprocal of $f(i)$, which is expressed as:

$$w_i = \frac{f(i)^{-1}}{\sqrt{f(0)^{-2} + f(1)^{-2} + \ldots + f(k-1)^{-2}}} \tag{16}$$

where the purpose of denominator in the above expression is to normalize the weighting coefficients vector to have a unit-norm. Based on the above setup, the last remaining step is to determine the optimum values of k and λ that minimize the reconstruction error, which is to be done numerically in Section IV-C.

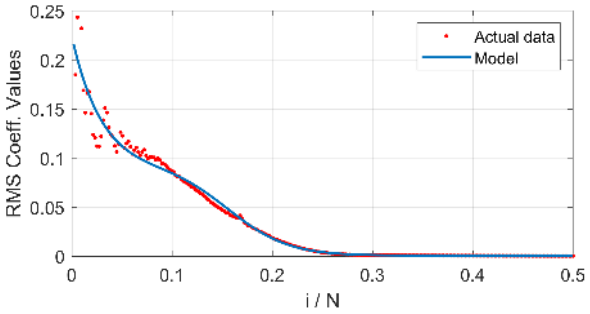


FIGURE 4: RMS values of DCT coefficients versus coefficient index computed MIT-BIH database records no. 100, 112, 123, 200 and 222 at $N = 512$. All frames were normalized to have a unit norm prior to calculation.

C. COMPUTATIONAL COMPLEXITY ANALYSIS

The computational complexity of FCE decoding can be calculated by analyzing (14). To simplify the analysis, we may recast (14) as determining the LS solution of $\mathbf{G}\mathbf{u}_k = \mathbf{b}$, where $\mathbf{G} = \mathbf{H}_k^T \mathbf{H}_k + \lambda \mathbf{W}_k^2$ is a symmetric and positive definite $k \times k$ matrix (since $\mathbf{W}_k$ is a diagonal matrix), and $\mathbf{b} = \mathbf{H}_k^T \mathbf{y}$ is a $k \times 1$ vector. The solution of this equation can be efficiently and stably found by using Cholesky decomposition method [25]. The main computational tasks associated with this process are listed as follows:

- Computing $\mathbf{H}_k = \mathbf{\Phi}\mathbf{\Psi}_k$ requires $\mathcal{O}(dkN)$ and $\mathcal{O}(kN)$ operations for RSBM and RD sensing matrices, respectively.
- Computing $\mathbf{G} = \mathbf{H}_k^T \mathbf{H}_k + \lambda \mathbf{W}_k^2$ requires $\mathcal{O}(k^2 M)$ operations, mainly due to the first term of the right-hand side.
- Computing $\mathbf{b} = \mathbf{H}_k^T \mathbf{y}$ requires $\mathcal{O}(kN)$ operations.
- Given $\mathbf{G}$ and $\mathbf{b}$, computing $\mathbf{u}_k$ by using Cholesky decomposition requires $\mathcal{O}(k^3)$ operations [25].

Asymptotically, $\mathcal{O}(k^2 M)$ dominates if $k < M$, while $\mathcal{O}(k^3)$ dominates if $k > M$.

On the other hand, the asymptotic complexity of solving the BPDN problem by using the SPG-L1 and BSBL algorithms is $\mathcal{O}(M^2 N)$ [31], [50]. WLM approach presents a slight modification to BPDN and has the same complexity [21]. On the other hand, the OMP and WOMP algorithms complexity is $\mathcal{O}(sMN)$ [11], [22]. Since $k < N$, we expect that FCE reconstruction should theoretically run faster than SPG-L1 and BSBL, and to be comparable in speed to OMP and WOMP. However, a major strength of FCE in this context is that while all CS reconstruction algorithms require multiple iterations to converge [11], [12], solving (14) is achieved in a single iteration, which makes it significantly faster. This will be verified by the numerical results presented in Section IV-E.

D. COMPARISON TO WEIGHTED ℓ_1 MINIMIZATION TECHNIQUE

At this point, it is necessary to to compare our proposed FCE technique to WLM [21], which bears some similarity in approach to FCE. The main observation that drove the WLM technique was the rapid decay of the DWT detail coefficients with increased resolution level. This can be observed by referring back to the DWT representation shown in Fig. 3. Consequently, an $N \times N$ diagonal weighting matrix $\mathbf{W} = \text{diag}(w_0, w_1, \ldots, w_{N-1})$ was plugged into the ℓ_1 minimization part of the BPDN objective function. This aims to incorporate the prior information about the coefficients decay over each resolution level. The WLM problem is given by the following form [21]:

$$\hat{\mathbf{u}} = \arg\min_{\mathbf{u}} \ \tfrac{1}{2}\|\mathbf{y} - \mathbf{H}\mathbf{u}\|_2^2 + \lambda \|\mathbf{W}\mathbf{u}\|_1 \tag{17}$$

The above formulation is an alternative form of the BPDN problem in (5), called the *lasso* problem [29].

By comparing the above problem to (13), we note that both FCE and WLM employ a diagonal weighting matrix to exploit prior knowledge of coefficients decay characteristics. However, we highlight the factors that distinguish FCE from WLM as follows:

1) *Solution characteristics*: WLM targets reconstructing a sparse DWT representation vector of length N, while FCE targets reconstructing a dense vector comprising the first k coefficients of the DCT representation.

2) *Weighting matrix and regularization function*: the WLM weighting matrix is designed to assign higher weights to significant coefficients, which leads ℓ_1 minimization to converge at a solution with minimum number of non-zeros. On the other hand, FCE weighting matrix is designed to assign low weights to significant coefficients. This leads the weighted optimum solution to have low energy. Hence, it can be approached via ℓ_2 norm minimization.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3035423, IEEE Access

3) *Degrees of freedom*: the performance of WLM is tuned solely by the parameter λ, while FCE is tuned by both λ and k. In addition, since k determines the size of the WRLS problem, selecting a smaller k contributes to increasing decoding speed.

4) *Path to solution*: any CS decoding algorithm used to solve the WLM problem requires multiple iterations to converge [11], [12], while FCE has a closed-form solution.

The above factors have lead FCE to achieve improvements over WLM in terms of reconstruction quality and speed. The same arguments apply to the WOMP technique [23], which is an adaptation of WLM to the OMP algorithm. We verify the advantages of FCE over WLM and WOMP and other decoding techniques through numerical experiments in the next section.

IV. RESULTS AND DISCUSSION

A. EXPERIMENTAL SETUP

ECG data used in all numerical experiments shown in this paper were extracted from the MIT-BIH arrhythmia database [47], [48]. The database includes a library of 30 minutes long ECG records of patients. The ECGs were sampled at a 360 Hz with 16-bit resolution, i.e. the length of each record is approximately 65×10^4 samples. The records are labeled 100 to 124 and 200 to 234. We arbitrarily selected records number 100, 112, 123, 200 and 222 for our experiments. The first 75% of each record data were used for selecting the optimum FCE parameters: k and λ in Section IV-C. The remaining 25% were used for producing decoding performance results in Section IV-D. This ensures unbiased evaluation of the decoding performance. On the other hand, the entire available data were used for measuring average execution time of CS decoding in Section IV-E.

Prior to CS encoding, the ECG signals were filtered by 4-pole Butterworth high-pass and low-pass filters. The filters' cutoff frequencies are 0.5 Hz and 40 Hz, respectively, as specified for ambulatory ECG monitoring [51], [52]. Due to their impact on minimizing the CS encoder complexity, we selected RSBM with $d = 12$ and RD sensing matrices for CS encoding in our experiments. The acquired CS measurements were quantized using 11 bits to better emulate realistic systems. For DWT basis, we used Symlet-6 wavelet with 6 levels of decomposition, following the setting used for WLM and WOMP in [21], [23]. Available ECG records were divided into frames of length N each. Presented results are the average of randomized trials performed on all available frames independently. Each trial uses a different random instance of the sensing matrix.

Numerical results for CS-ECG decoders included in our experiments were produced by the following publicly available MATLAB-based solvers: SPGL1 solver for BPDN [53], BSBL_BO solver for BSBL [54], l1-ls solver for the lasso formulation used in WLM [55], and OMP solver from Sparselab toolbox [56] for WOMP, with applying relevant modifications according to [23]. Finally, all numerical experiments were performed by using MATLAB 2018a running on a desktop computer equipped with an octa-core Intel i7-10700 processor operating at 2.9 GHz, and 16 GB of DDR4 RAM operating at 3.2 GHz.

B. RECONSTRUCTED SIGNAL QUALITY METRICS

The reconstruction quality in ECG compression literature is quantified by either the Percentage Root-mean-square Difference (PRD) or the Average Reconstruction Signal-to-Noise Ratio (ARSNR). The PRD is defined as [13], [57]:

$$\mathrm{PRD} \triangleq \sqrt{\mathbb{E}_{\mathbf{x}} \left(\frac{\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2}{\|\mathbf{x}\|_2^2} \right)} \times 100\% \tag{18}$$

The associated ARSNR in dB is defined as:

$$\mathrm{ARSNR} \triangleq 10 \log_{10} \mathbb{E}_{\mathbf{x}} \left(\frac{\|\mathbf{x}\|_2^2}{\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2} \right) \tag{19}$$

Based on the perception of a medical specialist, the relationship between the PRD/ARSNR and the visual quality of the reconstructed ECG signal is shown in Table 1 [58].

TABLE 1: Reconstructed ECG signal quality grades and the corresponding PRD and ARSNR values as perceived by a medical specialist [58].

Quality Grade	PRD	ARSNR
Very good (VG)	$0 \sim 2\%$	≥ 34 dB
Good (G)	$2 \sim 9\%$	$21 \sim 34$ dB
Indeterminable	$> 9\%$	< 21 dB

C. OPTIMUM FCE PARAMETERS SELECTION

In this section, we perform numerical experiments to determine optimal values of the parameters k and λ in (14) that maximize the ARSNR. Fig. 5 depicts the ARSNR versus k/N for $\lambda = \{0.01, 0.1, 1, 10\}$ at CR of 60% and 80%. We first notice that the ARSNR increases gradually with k, and then saturates at the peak achievable ARSNR. We also notice that very small values of λ ($\lambda = 0.01, 0.1$) lead to high performance variability, especially around $k = M$. This is referred to the fact that for $k = M$ and $\lambda \approx 0$, the solution in (14) will converge to $\hat{\mathbf{u}}_k \approx \mathbf{H}_k^{-1}\mathbf{y}$, i.e. this solution assumes $\mathbf{y} \approx \mathbf{H}_k\mathbf{u}_k$, which is inaccurate by referring to (10). On the other hand, a large value of $\lambda = 10$ leads to higher stability, since the regularizing term helps leading to the optimum solution, whilst peak ARSNR is reached at a larger value of k. We favor choosing a small k to minimize the problem size (and hence the computational task). Hence, our choice of the optimum value of k would be the smallest value that achieves the peak ARSNR. Therefore, we can deduce that an intermediate value $\lambda = 1$ is optimal.

Table 2 lists the optimum k values: k^* (normalized to N and M) that maximize the ARSNR for a set of CRs. We observe that for CR = 75% and above, $k^* > M$. This emphasizes the contribution of the regularizing energy minimization term in (13), which enabled improving recovery performance at high CR. Finally, we highlight that the values of k^* may vary for different system settings, such

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3035423, IEEE Access

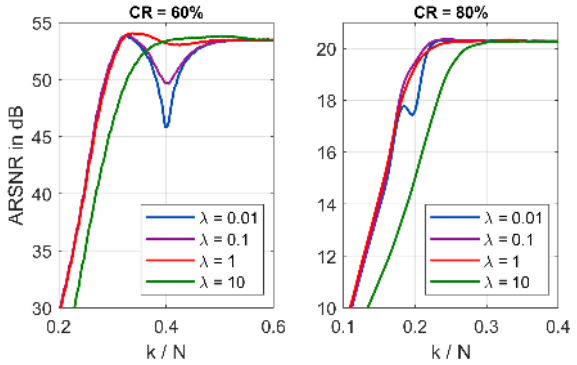


FIGURE 5: ARSNR versus k/N for CR = 60% and 80%. The results were produced by solving (14) for each value of k at $N = 512$ by using RSBM sensing matrix.

as the sampling rate, prefiltering cutoff frequencies and the measurements noise level. Hence, they should be fine-tuned by the system designer according to system specifications.

TABLE 2: A list of numerically calculated optimum values of the parameter k that maximize the ARSNR at $\lambda = 1$.

CR	k^*/N	k^*/M	CR	k^*/N	k^*/M
40%	0.35	0.58	70%	0.32	1.07
50%	0.35	0.70	75%	0.29	1.16
55%	0.34	0.76	80%	0.25	1.25
60%	0.33	0.83	85%	0.22	1.47
65%	0.32	0.91	90%	0.18	1.80

D. RECONSTRUCTED SIGNAL QUALITY

Fig. 6 depicts ARSNR against CR for FCE and a set of related CS-ECG decoding techniques. The results are shown for RSBM and RD sensing matrices. We first notice that the ARSNR associated with FCE is significantly higher than all other alternatives for both RSBM and RD sensing matrices. This is valid up to CR of 80%, after which WLM and WOMP achieve comparable performance for RSBM sensing matrix. However, at this CR range, the resulting ARSNR is too low to be practically useful. FCE performance is particularly superior for RD sensing matrix, where all tested CS-ECG decoding techniques achieved poor performance. This agrees with the discussion in Section II-B. We also notice that FCE achieves an ARSNR above 40 dB at CR $\leq 70\%$. As a result, FCE is particularly useful in operation modes that target high-quality reconstruction in return for lower CR.

Table 3 compares the maximum CR achievable by each technique for “Very Good" (VG) and “Good" (G) reconstructed signal quality grades, based on Table 1. The data show that FCE can achieve VG grade at CR up to 74% for RSBM sensing matrix, compared to 56% for WOMP, 54% for WLM, and 50% for BSBL. Consequently, for RSBM sensing matrix, FCE achieves VG grade with at least 32% higher CR than other techniques. On the other hand, none of the these techniques could achieve VG grade for RD sensing matrix, while FCE maintains its performance. Therefore, FCE has a good potential to perform efficiently over the RD

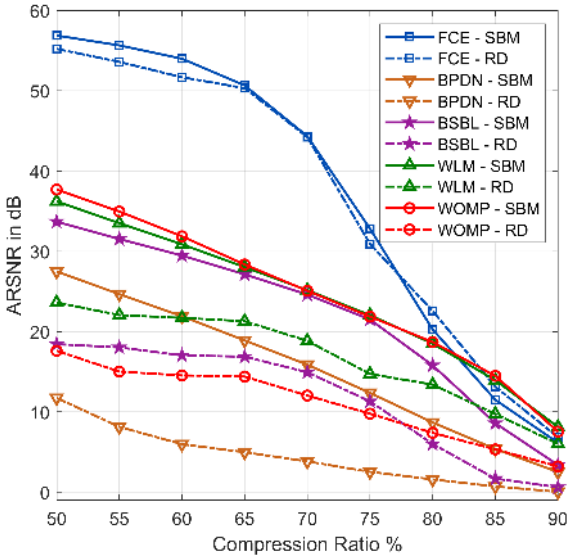


FIGURE 6: ARSNR versus CR at $N = 512$.

acquisition architecture at the sensing node. This grants a significant reduction in the energy consumption and cost of the CS encoder [41].

TABLE 3: Maximum achievable CR for VG and G quality grades.

	RSBM sensing matrix		RD sensing matrix	
Method	VG grade	G grade	VG grade	G grade
FCE	74%	80%	74%	81%
BPDN	N/A	62%	N/A	N/A
BSBL	50%	75%	N/A	N/A
WLM	54%	76%	N/A	66%
WOMP	56%	76%	N/A	N/A

In Fig. 7, we assess and compare the variability of the decoding performance by presenting a box plot of the PRD values achieved by all tested methods. The results shown are computed at CR = 75% and $N = 512$ for the entire record number 200 of the MIT-BIH database. This constitutes approximately 1270 ECG frames. For each box, the central line, lower and upper edge resemble the median, 25th and 75th percentile, receptively. The whiskers extend to the lower and upper extrema. Outliers are marked by the (+) symbol. Results for RSBM sensing matrix show that FCE has very low performance variability. FCE performance is followed by WLM, BSBL and WOMP, respectively. However, the latter three have higher median and a larger spread. On the other hand, for RD sensing matrix, WOMP performance is most impacted, while the impact on the performance of BSBL and WLM is less severe. However, FCE preserves its consistency. Finally, in all cases, BPDN performance is quite poor, as expected.

To visually demonstrate and compare the quality of the reconstructed ECG signal, Fig. 8 illustrates two ECG frames of 3 seconds each, which were taken from the MIT-BIH

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3035423, IEEE Access

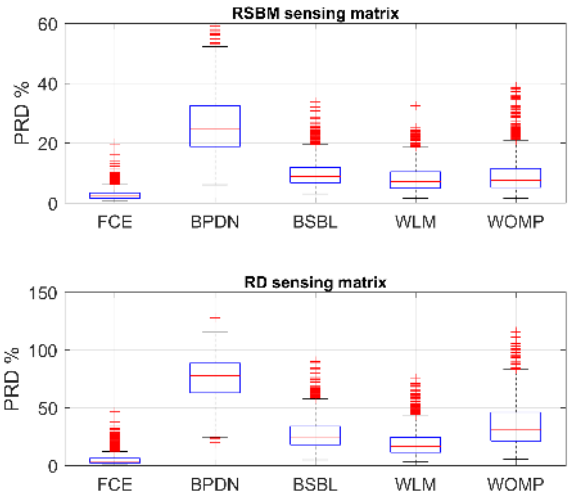


FIGURE 7: Box plots of the reconstruction PRD for MIT-BIH record no. 200 at CR = 75% and $N = 512$.

database. A normal ECG frame is shown in Fig. 8a and an abnormal frame is shown in Fig. 8b. The figure compares the original signals to those reconstructed by FCE, WOMP (with RSBM sensing matrix) and WLM (with RD sensing matrix) at CR = 75%. According to Fig. 6 and Fig. 7, the latter techniques achieved the second-best performance after FCE. We observe several distortions in the signals reconstructed by WOMP, while the QRS complex is highly affected in the signals reconstructed by WLM. On the other hand, for FCE, QRS complexes and other indicative intervals are well-defined and clear, and there is no noticeable distortion or other signal artifacts. Thus, it is clear that the reconstruction quality of FCE is superior.

E. DECODING SPEED

Fig. 9 depicts the average execution time for CS decoding versus $N = \{256, 512, 1024, 2048\}$ at CR = 75%. We observe that the FCE execution time is significantly shorter than all other tested CS-ECG techniques. According to the speed gains listed in Table 4, FCE is on average 33 times faster than WOMP, and several hundred times faster than the other methods. We may hence deduce that FCE reconstruction has an obvious speed advantage over other methods, especially for large values of N. In addition, storage space requirement of FCE is accordingly reduced, since the compression matrix size was shrunk from $M \times N$ to become $M \times k$. All temporary storage space needed for intermediate calculations is consequently shrunk.

V. CONCLUSIONS

In this paper, we presented a Fast Compressive Electrocardiography (FCE) technique. FCE is a low-complexity CS decoding technique that has been tailored for optimal reconstruction of compressed ECG signals. Numerical results demonstrated considerable improvements compared to other techniques presented in related literature, in terms of decoding speed and quality. FCE can contribute to minimizing the computational and energy cost of CS decoding, especially when decoding is done at the resource-constrained WBSN gateway node. Thus, enabling real-time decoding in long-term ECG monitoring scenarios. In addition, FCE is compatible with the random demodulator acquisition architecture, which enables energy-efficient and low-cost implementation of the CS encoder on the wearable ECG sensor module. Future extensions of this work include adapting FCE to deal with realistic implementation challenges, especially large CS measurements errors and invalid measurements caused by wireless channel impairments [59], [60]. In addition, developing a hardware implementation of an FCE-based ECG monitoring system within a smart-home framework [61]. Realistic ECG reconstruction performance and energy consumption are then measured, evaluated and compared.

TABLE 4: Execution speed gains of FCE against other tested CS-ECG decoding techniques.

N	BPDN	BSBL	WLM	WOMP
256	771	477	631	23
512	619	510	611	24
1024	258	465	383	28
2048	117	547	428	58
Average	441	500	513	33

ACKNOWLEDGMENTS

The author acknowledges Zhang Xin for her motivating and inspiring introduction to practical ECG monitoring instruments and procedures. The author also acknowledges Bilal Rateb for the curious discussions and the appreciated company throughout the whole process of writing this paper.

REFERENCES

[1] A. M. Nia, M. Mozaffari-Kermani, S. Sur-Kolay, A. Raghunathan, and N. K. Jha, "Energy-Efficient Long-term Continuous Personal Health Monitoring," IEEE Transactions on Multi-Scale Computing Systems, vol. 1, no. 2, pp. 85–98, apr 2015.

[2] K. Zhang and W. Ling, "Health Monitoring of Human Multiple Physiological Parameters Based on Wireless Remote Medical System," IEEE Access, vol. 8, pp. 71 146–71 159, 2020.

[3] M. Elgendi, A. Mohamed, and R. Ward, "Efficient ECG Compression and QRS Detection for E-Health Applications," Scientific Reports, vol. 7, no. 1, p. 459, dec 2017.

[4] T. Tekeste, H. Saleh, B. Mohammad, and M. Ismail, "Ultra-low power QRS detection and ECG compression architecture for IoT healthcare devices," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 66, no. 2, pp. 669–679, 2019.

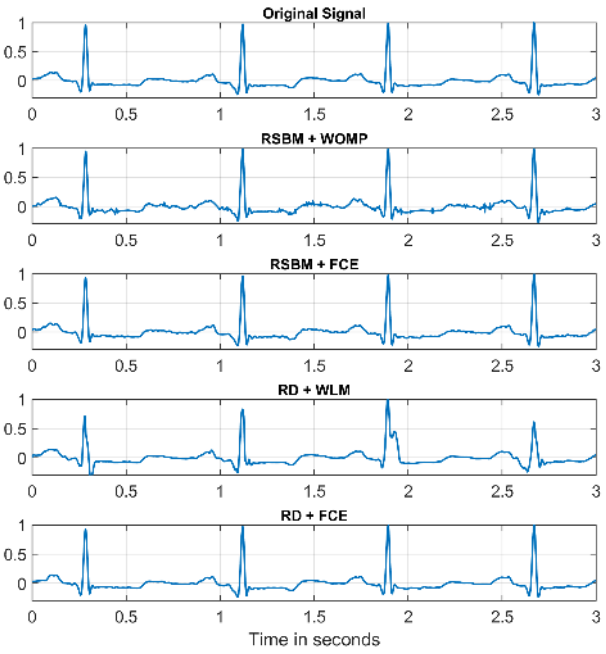
[5] D. Bortolotti, M. Mangia, A. Bartolini, R. Rovatti, G. Setti, and L. Benini, "Energy-Aware Bio-Signal Compressed Sensing Reconstruction on the WBSN-Gateway," IEEE Transactions on Emerging Topics in Computing, vol. 6, no. 3, pp. 370–381, jul 2018.

[6] M. Al Disi, H. Djelouat, C. Kotroni, E. Politis, A. Amira, F. Bensaali, G. Dimitrakopoulos, and G. Alinier, "ECG signal reconstruction on the IoT-gateway and efficacy of compressive sensing under real-time constraints," IEEE Access, vol. 6, pp. 69 130–69 140, 2018.

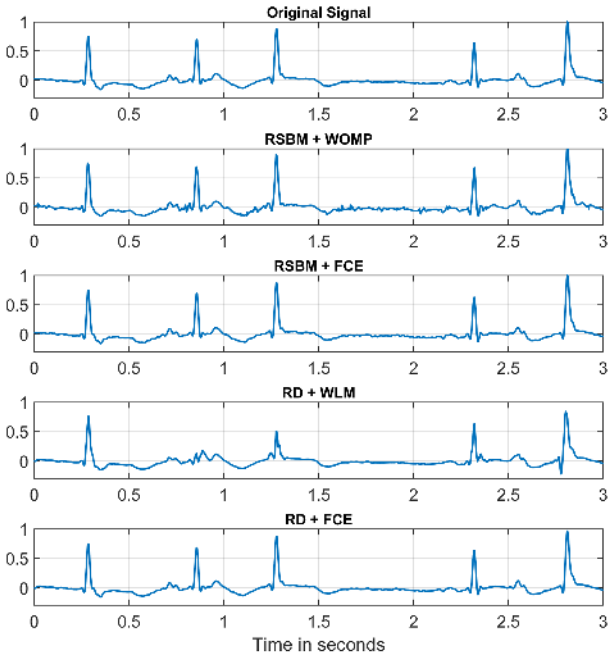
[7] P. K. Jain and A. K. Tiwari, "Heart monitoring systems—a review," Computers in Biology and Medicine, vol. 54, pp. 1 – 13, 2014.

[8] F. Sun, C. Kuo, and M. Griss, "PEAR: Power efficiency through activity recognition (for ecg-based sensing)," in 2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops, 2011, pp. 115–122.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3035423, IEEE Access



(a) Normal ECG frame taken from record no. 100.



(b) Abnormal ECG frame taken from record no. 222.

FIGURE 8: Visual comparison of a normal and an abnormal ECG frames and their reconstructed versionss at CR = 75%. Frame length is $N = 1080$, which is equivalent to 3 seconds of ECG readings. Amplitude is normalized to the maximum peak of the original signal.

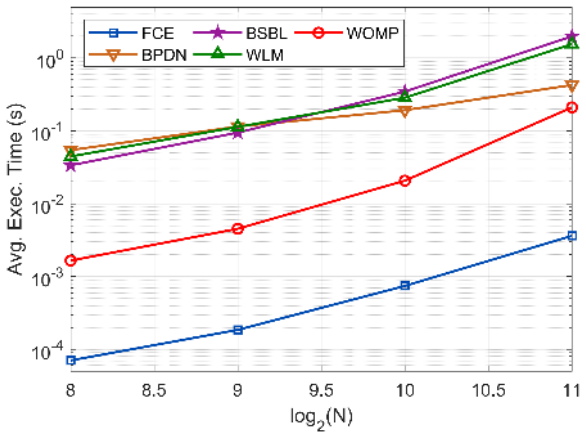


FIGURE 9: Average execution time for CS decoding versus $\log_2 N$ at CR = 75%.

[9] D. Donoho, "Compressed Sensing," IEEE Transactions on Information Theory, vol. 52, no. 4, pp. 1289–1306, apr 2006.

[10] J. Chen, J. Xing, L. Y. Zhang, and L. Qi, "Compressed sensing for electrocardiogram acquisition in wireless body sensor network: A comparative analysis," International Journal of Distributed Sensor Networks, vol. 15, no. 7, p. 1550147719864884, 2019.

[11] J. A. Tropp and S. J. Wright, "Computational Methods for Sparse Solution of Linear Inverse Problems," Proceedings of the IEEE, vol. 98, no. 6, pp. 948–958, jun 2010.

[12] E. Crespo Marques, N. Maciel, L. Naviner, H. Cai, and J. Yang, "A review of sparse recovery algorithms," IEEE Access, vol. 7, pp. 1300–1322, 2019.

[13] 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, Sep. 2011.

[14] D. H. Chae, Y. F. Alem, S. Durrani, and R. A. Kennedy, "Performance Study of Compressive Sampling for ECG Signal Compression in Noisy and Varying Sparsity Acquisition," in 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, may 2013, pp. 1306–1309.

[15] C. Karakus, A. C. Gurbuz, and B. Tavli, "Analysis of Energy Efficiency of Compressive Sensing in Wireless Sensor Networks," IEEE Sensors Journal, vol. 13, no. 5, pp. 1999–2008, may 2013.

[16] V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A Case Study in Low-Complexity ECG Signal Encoding: How Compressing is Compressed Sensing?" IEEE Signal Processing Letters, vol. 22, no. 10, pp. 1743–1747, oct 2015.

[17] M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "Adapted compressed sensing: A game worth playing," IEEE Circuits and Systems Magazine, vol. 20, no. 1, pp. 40–60, 2020.

[18] Z. Zhang, T. Jung, S. Makeig, and B. D. 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, Feb 2013.

[19] Z. Zhang and B. D. Rao, "Extension of sbl algorithms for the recovery of block sparse signals with intra-block correlation," IEEE Transactions on Signal Processing, vol. 61, no. 8, pp. 2009–2015, April 2013.

[20] L. F. Polanía and R. I. Plaza, "Compressed sensing ECG using restricted boltzmann machines," Biomedical Signal Processing and Control, vol. 45, pp. 237 – 245, 2018.

[21] J. Zhang, Z. Gu, Z. L. Yu, and Y. Li, "Energy-efficient ECG compression on wireless biosensors via minimal coherence sensing and weighted ℓ_1 minimization reconstruction," IEEE Journal of Biomedical and Health Informatics, vol. 19, no. 2, pp. 520–528, March 2015.

[22] J. A. Tropp and A. C. Gilbert, "Signal Recovery from Random Measurements via Orthogonal Matching Pursuit," IEEE Transactions on Information Theory, vol. 53, no. 12, pp. 4655–4666, 2007.

[23] A. Marchioni, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "Low-complexity greedy algorithm in compressed sensing for the adapted decoding of ecgs," in 2017 IEEE Biomedical Circuits and Systems Conference (BioCAS), 2017, pp. 1–4.

[24] F. Pareschi, M. Mangia, D. Bortolotti, A. Bartolini, L. Benini, R. Rovatti, and G. Setti, "Energy analysis of decoders for rakeness-based compressed sensing of ecg signals," IEEE Transactions on Biomedical Circuits and Systems, vol. 11, no. 6, pp. 1278–1289, 2017.

[25] S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge University Press, 2004.

[26] A. V. Oppenheim and R. W. Schafer, Discrete-Time Signal Processing, 3rd ed. USA: Prentice Hall Press, 2009.

[27] E. Correia Pinheiro, O. A. Postolache, and P. Silva Girão, "Implementation of Compressed Sensing in Telecardiology Sensor Networks," International Journal of Telemedicine and Applications, vol. 2010, pp. 1–12, 2010.

[28] M. S. Manikandan and S. Dandapat, "Wavelet-based electrocardiogram signal compression methods and their performances: A prospective review," Biomedical Signal Processing and Control, vol. 14, pp. 73 – 107, 2014.

[29] J. A. Tropp, "Just Relax: Convex Programming Methods for Identifying Sparse Signals in Noise," IEEE Transactions on Information Theory, vol. 52, no. 3, pp. 1030–1051, 2006.

[30] S. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale ℓ_1-regularized least squares," IEEE Journal of Selected Topics in Signal Processing, vol. 1, no. 4, pp. 606–617, 2007.

[31] E. G. Birgin, J. M. Martínez, and M. Raydan, "Spectral Projected Gradient Methods: Review and Perspectives," Journal of Statistical Software, vol. 60, no. 3, pp. 1–21, 2014.

[32] 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, may 2009.

[33] E. J. Candes, J. K. Romberg, and T. Tao, "Stable Signal Recovery from Incomplete and Inaccurate Measurements," Communications on Pure and Applied Mathematics, vol. 59, no. 8, pp. 1207–1223, 2006.

[34] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A Simple Proof of the Restricted Isometry Property for Random Matrices," Constructive Approximation, vol. 28, no. 3, pp. 253–263, jan 2008.

[35] A. M. Rateb and S. K. Syed-Yusof, "Performance analysis of compressed sensing given insufficient random measurements," ETRI Journal, vol. 35, no. 2, pp. 200–206, 2013.

[36] A. M. Rateb and S. K. S. Yusof, "Recovery error bounds on compressed sensing of noisy signals," International Journal of Communication Systems, vol. 28, no. 3, pp. 546–559, 2015.

[37] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," IEEE Transactions on Information Theory, vol. 47, no. 7, pp. 2845–2862, 2001.

[38] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," IEEE Transactions on Information Theory, vol. 49, no. 12, pp. 3320–3325, 2003.

[39] Z. Ben-Haim, Y. C. Eldar, and M. Elad, "Coherence-based performance guarantees for estimating a sparse vector under random noise," IEEE Transactions on Signal Processing, vol. 58, no. 10, pp. 5030–5043, 2010.

[40] S. Kirolos, J. Laska, M. Wakin, M. Duarte, D. Baron, T. Ragheb, Y. Massoud, and R. Baraniuk, "Analog-to-information conversion via random demodulation," in 2006 IEEE Dallas/CAS Workshop on Design, Applications, Integration and Software, 2006, pp. 71–74.

[41] J. N. Laska, S. Kirolos, M. F. Duarte, T. S. Ragheb, R. G. Baraniuk, and Y. Massoud, "Theory and implementation of an analog-to-information converter using random demodulation," in 2007 IEEE International Symposium on Circuits and Systems, 2007, pp. 1959–1962.

[42] J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk, "Beyond nyquist: Efficient sampling of sparse bandlimited signals," IEEE Transactions on Information Theory, vol. 56, no. 1, pp. 520–544, 2010.

[43] J. Yoo, S. Becker, M. Monge, M. Loh, E. Candès, and A. Emami-Neyestanak, "Design and implementation of a fully integrated compressed-sensing signal acquisition system," in 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012, pp. 5325–5328.

[44] F. Pareschi, P. Albertini, G. Frattini, M. Mangia, R. Rovatti, and G. Setti, "Hardware-algorithms co-design and implementation of an analog-to-information converter for biosignals based on compressed sensing," IEEE Transactions on Biomedical Circuits and Systems, vol. 10, no. 1, pp. 149–162, Feb 2016.

[45] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vandergheynst, "Design and exploration of low-power analog to information conversion based on compressed sensing," IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 2, no. 3, pp. 493–501, 2012.

[46] D. Gangopadhyay, E. G. Allstot, A. M. R. Dixon, K. Natarajan, S. Gupta, and D. J. Allstot, "Compressed sensing analog front-end for bio-sensor applications," IEEE Journal of Solid-State Circuits, vol. 49, no. 2, pp. 426–438, Feb 2014.

[47] G. B. Moody and R. G. Mark, "The impact of the mit-bih arrhythmia database," IEEE Engineering in Medicine and Biology Magazine, vol. 20, no. 3, pp. 45–50, 2001.

[48] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng, and H. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. circulation [online]," p. e215–e220, 2000.

[49] K. R. Rao and P. Yip, Discrete Cosine Transform. Academic Press, 1990.

[50] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse bayesian learning," IEEE Journal of Selected Topics in Signal Processing, vol. 5, no. 5, pp. 912–926, Sep. 2011.

[51] K. Venkatachalam, J. E. Herbrandson, and S. J. Asirvatham, "Signals and signal processing for the electrophysiologist. part i: Electrogram acquisition," Circulation: Arrhythmia and Electrophysiology, vol. 4, no. 6, pp. 965–973, 2011.

[52] D. Ricciardi et al., "Impact of the high-frequency cutoff of bandpass filtering on ecg quality and clinical interpretation: A comparison between 40hz and 150hz cutoff in a surgical preoperative adult outpatient population," Journal of Electrocardiology, vol. 49, no. 5, pp. 691 – 695, 2016.

[53] E. V. D. Berg and M. P. Friedlander, "SPGL1: A solver for large-scale sparse reconstruction," December 2019. [Online]. Available: https://friedlander.io/spgl1

[54] Z. Zhang, "BSBL for Recovery of Block Sparse Signal with Known or Unknown Block Partition." [Online]. Available: http://dsp.ucsd.edu/ zhilin/BSBL.html

[55] K. Koh, S.-J. Kim, and S. Boyd, "l1_ls: Simple Matlab Solver for l1-regularized Least Squares Problems." [Online]. Available: https://web.stanford.edu/ boyd/l1_ls/

[56] V. Stodden et al., "SparseLab: Seeking Sparse Solutions to Linear Systems of Equations," 2010. [Online]. Available: http://sparselab.stanford.edu/

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

[58] 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.

[59] M. Cheffena, "Performance Evaluation of Wireless Body Sensors in the Presence of Slow and Fast Fading Effects," IEEE Sensors Journal, vol. 15, no. 10, pp. 5518–5526, oct 2015.

[60] A. M. Rateb, S. Syed-Yusof, and R. A. Rashid, "On the Impact of Prefiltering on Compressed Sensing in Presence of Invalid Measurements," IEEE Signal Processing Letters, vol. 24, no. 12, pp. 1886–1890, Dec 2017.

[61] M. R. Abd Rahim, R. A. Rashid, A. M. Rateb, M. A. Sarijari, A. S. Abdullah, A. H. F. A. Hamid, H. Sayuti, and N. Fisal, Service-Oriented Architecture for IoT Home Area Networking in 5G. John Wiley & Sons, Ltd, 2018, ch. 16, pp. 577–602.