22nd IMEKO TC4 International Symposium & 20th International Workshop on ADC Modelling and Testing
SUPPORTING WORLD DEVELOPMENT THROUGH ELECTRICAL&ELECTRONIC MEASUREMENTS
IASI, ROMANIA, September 14-15, 2017

Compression Study of Continuous-Time Sampled ECG Data for e-Health Applications

Mariam TLILI[1, 2], Manel BEN ROMDHANE[1], Asma MAALEJ[1], François RIVET[2], Dominique DALLET[2], Chiheb REBAI[1]

[1]GRESCOM Research Lab., SUP'COM, University of Carthage, Tunisia.manel.benromdhane@supcom.tn
[2]IMS Research Lab., University of Bordeaux, Bordeaux INP ENSEIRB-MATMECA, France.dominique.dallet@ims-bordeaux.fr

***Abstract* – The study presented in this paper describes the compressibility of continuous-time sampled electrocardiogram (ECG) signals. The compression method is based on the orthogonal matching pursuit (OMP) algorithm applied to discrete Haar, Daubechie and Biorthogonal wavelets. Comparative results show that Biorthogonal and Daubechie wavelet families offer up to 49 % and 43 % reductions of the overall number of amplitude samples, respectively. However, Haar wavelets, the less complex wavelet family, discard only 12 % of the data and lead to higher computation time than Daubechie family.**

Keywords – E-Health, ECG, LC-ADC, wavelet families, compression, continuous-time sampling.

I. INTRODUCTION

The electrocardiogram (ECG) signals have an important role in heart diseases diagnosis. Hence, ECG acquisition is a common medical practice in ambulatory and external heart monitoring applications[1-2]. In the recent years, with the advance of information and telecommunication technologies, the number of electrodes in a device as well as monitoring duration considerably increased. Thus, the vast amount of acquired data increasesthe devices constraints such as size, memory resources and power consumption [3].

As an example of a common ECG acquisition, the monitoring of cryptogenic strokes, using a 12-lead configuration over up to 24 hours with a sampling frequency ranging between 8 kHz and 320 kHz, produces up to 480 kilo bytes of samples per minute [4]. As a solution to relax the acquisition constraints on the device, digital data compression has been employed since the early age of ECG processing. In this context, transform methods like the discrete wavelet transform (DWT) are widely used to eliminate samples by encoding the sparse transform of the ECG data [5-6].

Thus far, most of the attempts to compress ECG data are interested in reducing data size independently from the acquisition step [5,7]. These attempts show good compression performances. However, when power consumption is a limitation, the conventional scheme of acquiring redundant data which is discarded after compression is not power efficient[1].

As ECG signals are sparse in time domain, signal dependent ADC architecture is appealing [6;2,7]. To save power consumption and to reduce sampling rates, the level-crossing ADC (LC-ADC) is proposed[8-9]. The converter only acquires the relevant information and ignores the non-significant variations of the signal.

Certainly, the LC-ADC reduces the analogfront-end's activity and power consumption. However, the need for ECG compression still exists. Firstly, this is due to the low compression capabilities of the LC-ADC in particular zones of the ECG which will be further detailed in this paper. Secondly, considering the critical resources in e-Health devices, it is interesting to achieve higher compression ratios (CR). Thus, the system would benefit from less power consumption and storage space than that reached by only the LC-ADC. For these reasons, the authors propose to study the compressibility of continuous-time sampled ECG at the LC-ADC output. To the best of the authors' knowledge, while many works have been interested in compression of discrete-time sampled data [5-6], no study has been reported about compressibility of continuous-time sampled ECG data namely with DWT.

The remainder of this paper is organized as follows. Section II presents the LC-ADC and its compression performances. Section III explains the methods used in this work including a brief description of the wavelet transform and the compression study steps. In Section IV, comparative performance results are discussed to identify the best wavelet type to compress the digital data. Finally, section V concludes and draws some future works.

II. LC-ADC FOR ECG COMPRESSION

The level-crossing (LC) sampling scheme has been

introduced since the 80's as a compression technique associated with digital encoding[10]. It has been studied with speech and biological signals [9,11].

A. LC-ADC architecture

The converter's block diagram is shown in Fig. 1. The input signal is continuously compared to two reference voltages. When a LC event is detected by the continuous-time comparison block, digital levels are updated. The digital-to-analog conversion (DAC) block produces the new analog reference voltages. Moreover, a control logic is used to set the digital value of the crossed level, ECG_out, as the output sample[8].

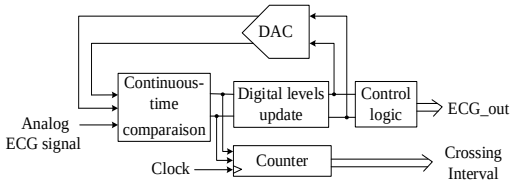

Fig. 1. Signal dependent LC-ADC block diagram.

For reconstruction of continuous-time sampled data, time information is needed to localize the LCs. Therefore, a counter computes the duration, $Crossing\ Interval$, between two consecutive LCs. An example of output data versus the original signal is given in Fig. 2.

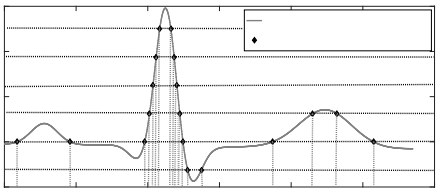

Fig. 2. Level-crossing sampling.

To measure LC-ADC performances in terms of data compression compared to discrete-time sampling, a detailed LC-ADC model has been implemented in Matlab/Simulink[8]. The next subsection presents some compression performances of the LC-ADC.

B. LC-ADC compression performances

The behavioral simulations are done with real ECG records from Physionet database. LC-ADC performances are measured in two cases: normal and myocardial infarction, a common and important medical emergency[12]. Thus, the tested records are 1-lead normal ECG and 15-lead myocardial infarction.

Firstly, the signals have been resampled at 1 MHz to model an analog signal. Secondly, since the converter's resolution, M, has an impact on the quantization step, the converter's model has been simulated with different values of M until a good quality reconstructed signal has been obtained[8]. Lastly, the CR has been computed between the converter's output and the original input according to (1)where $bit_{orig}$is the total bits of the original signal as indicated in Physionet and $bit_{LCADC}$is the total bits of LC-ADC samples.

$$CR(\%) = \frac{bit_{orig} - bit_{LCADC}}{bit_{orig}} * 100 \qquad (1)$$

The quantity of output bits includes Crossing Interval and $\mathrm{ECG_{out}}$values. Simulations have been carried out over a single period of each ECG which contains a P wave followed by a QRS complex then a T wave. The results are described in Table 1.

Table 1. CRs versus ECG signals and LC-ADC bit resolutions.

	M	CR (%)
1-lead Normal ECG (12 bits*)	9	35
	10	No compression
15-lead myocardial infarction (16 bits*)	9	33 to 63
	10	17 to 29

(*): number of bits of the original signal

As given in Table 1, 10-bit LC-ADC CRs are low. In the worst case, the total number of samples is higher than for a conventional discrete-time ADC and no compression is achieved. This is due to the signal dependent behavior of the converter which also depends on the distribution density of the reference levels. Thus, in high slope regions like the QRS complex, when the bit resolutionis increased, more level-crossings are detected which reduces the CR. This fact is also illustrated in Fig. 2 where 66.6 % of the output samples are representative of the QRS complex for 3-bit resolution. Therefore, compression is needed after the LC-ADC. Since wavelet methods are common in ECG compression, it is intuitive to inspect their performances when applied to continuous-time sampled data [13].

III. FROM DWT TO ECG COMPRESSION METHODOLOGY

In this paper, the compression study is done on signals that are sampled with respect to LC principle. The orthogonal matching pursuit (OMP) algorithm is used as the compression algorithm based on iterative application of DWT[14]. The computed output of the OMP is known as the K-sparse representation which contains the K most significant wavelets.K is provided as an input parameter to the algorithm and also defined as the maximum number of iterations since one wavelet is selected in every OMP iteration. In this section, before the DWT test, it is necessarily to present the wavelets. On the one hand, the authors

describe the DWT. On the other hand, they present the most used wavelets for discrete-time ECG signal compression. Finally, the test methodology of DWT-based ECG compression is announced.

A. Discrete wavelet transform for compression

The DWT is a set of linear operations that generates a vector of discrete data[14]. These operations are seen as projections of the original signal to a different domain which is described by the wavelet functions. The matrix form of these projections is given in (2)where x is the column of samples of the original signal, s is the column ofthe wavelet coefficients,α_i, and Ψ is an N-by-N matrix with wavelets, ψ_i,in column number i.

$$x\Psi = s \text{ with } \langle x, \psi_i \rangle = \alpha_i \tag{2}$$

It should be noted that, an active wavelet results in a high $\alpha_i$value. When x is compressible in Ψ, $s$is sparse meaning that the number of the most significant coefficients in s is small regarding the length of s. The recovery of x from these coefficients is done by application of the inverse of Ψ, Ψ^{-1}, to both sides of equation (2) which leads to equation (3).

$$x = s\Psi^{-1} \tag{3}$$

Using an orthogonal basis Ψ is computationally effective when it comes to signal recovery. When x is compressible in Ψ, and when Ψ is orthogonal, the recovery of the original signal can be from its sparse projections without high distortion as given in (4):

$$\hat{x} = \sum_{u=1}^{K} \alpha_u \psi_u \, , K \ll N \tag{4}$$

where $\hat{x}, K$ and N are the compressed version of x, the number of significant wavelets and the length of x.

B. Wavelet bases and multiresolution analysis

A wavelet family is obtained by translation and dilation of a mother wavelet, ψ_1, The orthogonal set of these functions, given in (5), is obtained by binary dilation and dyadic translation of the mother waveletwhere ψ_i, L, j and k are the i^{th} wavelet function,the decomposition level, the scaling factor and the translation factor, respectively.

$$\begin{cases} \psi_i(t) = 2^{\frac{j}{2}} \psi_1(2^j t - k) \\ i = 2^j + k, 0 < j \leq L, 0 \leq k < 2^j \end{cases} \tag{5}$$

In multiresolution analysis, it is possible to compute multilevel decompositions. In fact, the space spanned by Ψ is defined as the orthogonal sum of subspaces. Each subspace isspanned by the orthogonal basis, $\Psi^j, j \in \mathrm{Z}$. This basis is formed by the translated functions of a scaled version of the mother wavelet for a fixed value of j as described in (5) by varying the translation factor k, [14-15]. In this paper, the decomposition of digital data up to level L, refers to the projection of data to the subspaces spanned by$\{\Psi^j, j = 0 < j \leq L\} = \{\psi_i(t), 2j - 1 \leq i < 2^j, 0 \leq k < 2^j\}$ where $\psi_i(t)$ are the columns of Ψ^j. Thus, the multiresolution projection of x by selecting the K most significant wavelets of the set $\{k_j, 0 < j \leq L\}$ is described by (6)where k_j are addresses of the most significant wavelets inΨ^j.

$$\hat{x} = \sum_{j=1}^{L} \sum_{u \in \{k_j\}} \alpha_u \psi_u \tag{6}$$

In this paper, the authors choose three different wavelet families: Haar, Daubechie and Biorthogonal wavelets. On the one hand, Haar wavelets are selected due to their low complexity. On the other hand, the Daubechie and the Biorthogonal are widely used in conventional ECG data compression. Thus, it is interesting to investigate the compressibility of the continuous-time sampled ECG data in thesewavelets.

C. Haar, Daubechie and Biorthogonal wavelet families

Haar functions[16]are the simplest examples of wavelets. The mother wavelet is the piece wise function described in (7) whereas the wavelets obtained by translation and dilation of $\psi_1$are given in (8).

$$\psi_1(t) = \begin{cases} 1, & 0 \leq t \leq 1/2 \\ -1, & 1/2 \leq t \leq 1 \\ 0, & otherwise \end{cases} \tag{7}$$

$$\psi_i(t) = \begin{cases} 1, & \frac{k}{2^j} \leq t \leq \frac{k+1/2}{2^j} \\ -1, & \frac{k+1/2}{2^j} \leq t \leq \frac{k+1}{2^j} \\ 0, & otherwise \end{cases} \tag{8}$$

It should be noted that Haar basis as described in (8) is only invertible. The orthogonal form is described by other functions[15]. In this paper, the authors choose to implement the non-orthogonal form of Haar basis for its low complexity for future implementation.

As far as it concerns Daubechie wavelets, they are known for their compactness and orthogonality but they do not have a closed-form expression. To compute Daubechie wavelets, there are several numerical methods such as polynomial or matrix based methods[15,17]. However, in this paper, the authors use the predefined Maltab functions to generate the orthogonal matrix of Daubechie basis. Similarly, Biorthogonal wavelets are computed with the help of numerical algorithms. In this paper, the authors use the

translations and dilations of a B-spline function constructed by Matlab. For more extensive theory of wavelets, the authors refer the reader to the works of Daubechie and Cohen [14,18].

D. Methodology of ECG compression study

In this paper, x in (2) is replaced by ECG which is the LC-ADC output and $\hat{x}$in (4) is replaced by $\widehat{ECG}$.The block diagram of the compression study is presented in Fig. 3. Firstly, the signal from Physionet data base is up-sampled at 1 MHz to model. Secondly, the continuous-time samples are extracted thanks to LC-ADCmodel that is implemented in Matlab/Simulink[8]. Thirdly, in order to allow the multiplication between the D-by-D basis matrix, Ψ, and the converted data, the output vector's length is truncated to the nearest multiple of D using equation (9)where N and n are the original length of the digital data and the nearest multiple of D, respectively.

$$\widehat{N} = nD \text{ with } n = \left\lfloor \frac{N}{D} \right\rfloor \tag{9}$$

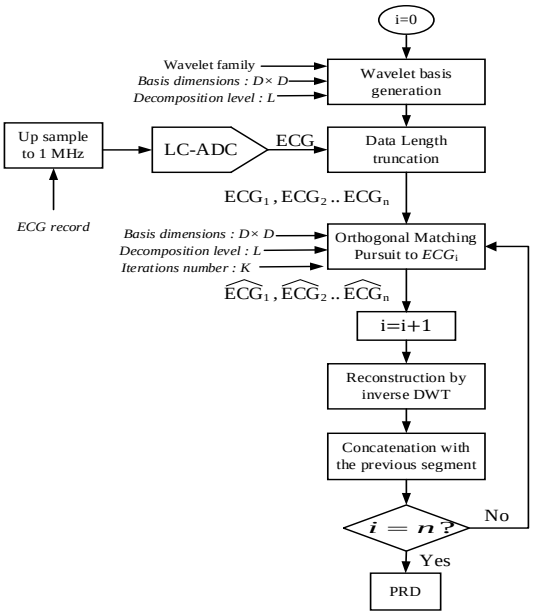

Fig. 3. Block diagram of ECG compression study.

After truncation, the vector of length nD is split into n segments, $ECG_i, i = 1..n$. Fourthly, the OMP is applied separately to each segment for a fixed value of the decomposition level, L, and for the maximum number of iterations, K. Fifthly, to recover the original signal, the inverse matrix of the wavelet basis is also applied separately to each ECG segment's decomposition. Once all the segments are recovered, the compressed ECG is found by concatenation of the recovered segments which leads to a signal of length nD with the nK most significant samples. Lastly, to evaluate the compression performances for fixed values of D, K and L, the percentage root difference (PRD) is computed for the nD length original and concatenated signals as described in (10)[19]where $ECG(i)$is the i^{th} sample of the truncated LC-ADC output, $\widehat{ECG}(i)$is the i^{th} sample of the recovered signal and $\overline{ECG}$is the mean value of the truncated LC-ADC output samples.

$$PRD = \sqrt{\frac{\sum_{i=1}^{nD}\left(ECG(i)-\widehat{ECG}(i)\right)^2}{\sum_{i=1}^{nD}(ECG(i)-\overline{ECG})^2}} \tag{10}$$

A 0 % PRDrepresents a lossless compression. Otherwise, it is important to keep the PRD lower than 2 % as the criteria for a very low distortion[19].

IV. PERFORMANCE RESULTS AND DISCUSSION

A. Simulation environment

A number of 67 different signals has been selected to conduct the study, namely: a normal ECG, two Gaussian models of a normal and a pathologic ECG [20], 15 leads of anterior myocardial infarction, 15 leads of inferior myocardial infarction and 43 other pathological signals. The design parameters of the Matlab/Simulink model are: 8-bit resolution, 10-kHz counter frequency and 12-bit counter resolution. Basis functions are generated using Matlab functions except Haar basis which is implemented. The advantage of the implemented version is the use of only ±1 in a sparse matrix. The OMP is applied with two decomposition levels as presented in Fig. 3.The simulation parameters are: wavelet family type, dimensions of the wavelet matrix, D, decomposition level, L, and the number K of retained functions $\psi_i$to represent the signal. For every signal, many simulations are done regarding$K$. The final step is the recovery of the original signal with the inverse of Haar and Biorthogonal bases matrices and the transpose of Daubechie matrix.

B. Simulation results and discussion

Simulation results are shown in Fig. 4 for Haar basis, in Fig. 5 for Daubechie basis and in Fig. 6 for Biorthogonal basis. The PRD decreases when the number of selected wavelets,K, increases. In fact, the more wavelets are retained to represent the signal, the less information is discarded. In the particular case of Haar wavelet family, with almost all tested signals, more than 88 % of the wavelet functions are needed to reach the 2 % PRD. This is true for level 2 and level 3 decompositions for 64 and 128 bases dimensions.When K is equal to 60 and 120, in the case of a 64 and 128 bases dimensions, respectively, allthe signals can be recovered with a PRD lower than2 %. In this case,a low PRD value cannot be reached without using atleast 93 % of the wavelets in the decomposition.

As far as it concerns the Daubechie wavelet family,

a 2 % PRD requires about 64 % and 57% of thewavelet functions for 64 and 128 dimensions, respectively, as shown in Fig. 5. These results are applicable to decomposition levels 2 and 3.Furthermore, referring to Fig. 6, Biorthogonal wavelet family gives the highest decompositionperformances.In fact, the 2 % PRD is reached when 55 % ofBiorthogonal wavelet functions are retained for 64 basis dimension.This value is reduced to 54 % and 51 % forlevels 2 and 3, respectively, if basis dimension is increased to 128.Moreover, it is seen in all simulations that the decomposition level has less effect than basis dimensions on PRD performances.

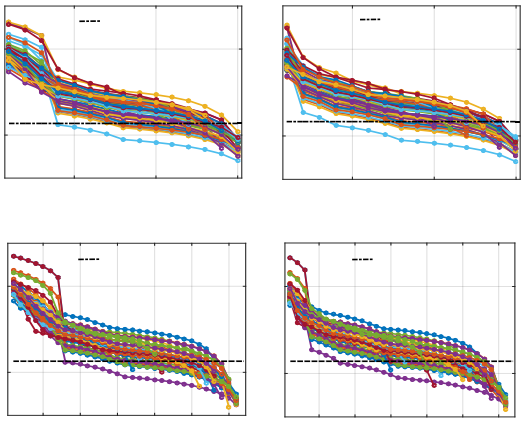

Fig. 4. PRD performances versus the number of wavelets, K, retained by the OMP for Haar basis.

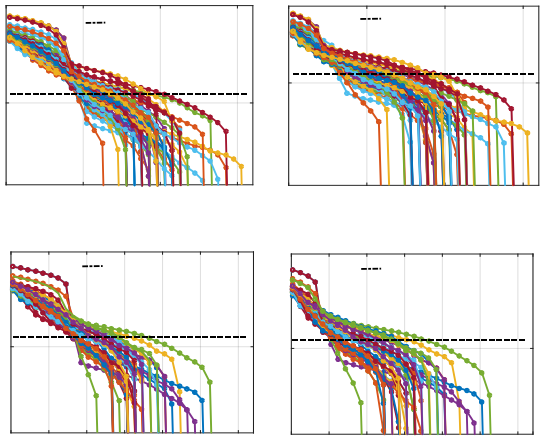

Fig. 5. PRD performances versus the number of wavelets, K, retained by the OMP for Daubechie basis

To sum up, simulation results confirm the non-efficiency of Haar basis to compress the continuous-time sampled ECG signals. In fact, low savings are done to eliminate the related coefficients.Contrariwise, the best *PRD* results for good CRs, around 50 %, are obtained with Biorthogonal wavelets decomposition using a 128 dimensions matrix. In fact, increasing the basis dimensions reduces the probability ofeliminating a high weighted wavelet function when$K$ varies. However, for LC-ADC output compression, these performances remain poor if the time is considered.

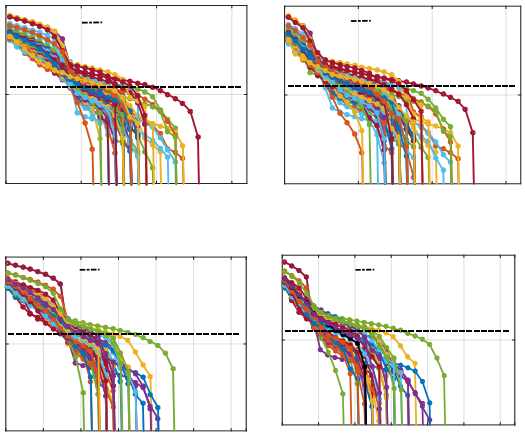

Fig. 6. PRD performances versus the number of wavelets, K, retained by the OMP for Biorthogonal basis.

Moreover, ECG data compression is inspected regarding occurrence probability of the wavelets in the signals decompositions. The less sparse signals are, the higher probability of occurrence is. In Fig. 7,the authors give simulation results of these probabilities values using the following parameters: $L = 2$ and $D = 64$ for Haar, Daubechie and Biorthogonal wavelets.

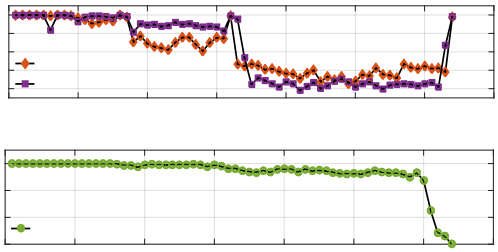

Fig. 7. Occurrence probability of wavelets in the ECG data decompositions for Haar, Daubechie and Biorthogonal bases.

Fig. 7 shows a probability of occurrence higher than 80 % of 60 Haar wavelets from the 64 basis functions. This result explains the high degradation of the recovered signals for low values of K. Thus, almost 91 % ofthe basis functions cannot be sidestepped. In contrast, a slighter improvement is seen with Daubechie and Biorthogonal families with probabilities higher than 60 % only for a number around 36 of the wavelet functions.However, it is noticed that wavelets 63 and 64 have to be also selected.

As a comparison regarding computation time, Table 2 summarizes simulation results in the case of one-lead

normal ECG leading to a 2 % PRD. The wavelets use2-level decomposition in a 64-dimension basis. It is seen that 54 Haar wavelets are needed to represent the signal. Consequently, the computation time is higher than Daubechie with only 23 wavelets. However, Biorthogonal basis requires the highest computational time. Therefore, Daubechie family gives the best compromise between complexity and data reduction.

Table2. Performance results: PRD = 2 %, D = 64 and L = 2.

	K	Computation time (ms)
Haar	54	31.4
Daubechie	23	20.7
Biorthogonal	21	76.6

V. CONCLUSIONS

In this paper, the authors study the compressibility of continuous-time sampled ECG data using threewavelet families. The study shows that Haar wavelets allow a 2 % PRD when at least 88 % of the basis wavelets represent a signal. Better results are observed with Biorthogonal and Daubechie wavelets which guarantee a PRD lower than 2 % for decompositions containing less than 64 % of the wavelets. However, DWT based methods are not sufficient when time information is considered. In this context, future works concern the proposal of a compression method that gives a low *PRD* for high CRs.

VI. ACKNOWLEDGMENT

This research work is part of the Wireless Biomedical Acquisition And Transmission (Wibio'ACT) project.The authors would like to thank the Comité Mixte de Coopération Universitaire for financial support of this project.

REFERENCES

[1] M. Elgendi, A. Mohamed, and R. Ward, "Efficient ECG Compression and QRS Detection for E-Health Applications," *Scientific Reports*, vol. 7, 2017.

[2] J. Wang, M. She, S. Nahavandi, and A. Kouzani, "Human identification from ECG signals via sparse representation of local segments," *IEEE Signal Processing Letters*, vol. 20, no. 10, pp. 937–940, 2013.

[3] M. Patel and J. Wang, "Applications, challenges, and prospective in emerging body area networking technologies," *IEEE Wireless communications*, vol. 17, no. 1, 2010.

[4] G. W. Albers, R. A. Bernstein, J. Brachmann, J. Camm, J. D. Easton, P. Fromm, S. Goto, C. B. Granger, S. H. Hohnloser, E. Hylek, and others, "Heart rhythm monitoring strategies for cryptogenic stroke: 2015 Diagnostics and Monitoring Stroke Focus Group Report," *Journal of the American Heart Association*, vol. 5, no. 3, p. e002944, 2016.

[5] X. Wang, Z. Chen, J. Luo, J. Meng, and Y. Xu, "ECG compression based on combining of EMD and wavelet transform," *Electronics Letters*, vol. 52, no. 19, pp. 1588–1590, 2016.

[6] U. Satija, B. Ramkumar, and M. S. Manikandan, "A unified sparse signal decomposition and reconstruction framework for elimination of muscle artifacts from ECG signal," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, 2016, pp. 779–783.

[7] G. Da Poian, R. Bernardini, and R. Rinaldo, "Sparse representation for fetal QRS detection in abdominal ECG recordings," in *E-Health and Bioengineering Conference (EHB), 2015*, 2015, pp. 1–4.

[8] M. Tlili, A. Maalej, M. Ben Romdhane, M. C. Bali, F. Rivet, D. Dallet, and C. Rebai, "Level-Crossing ADC Modeling for Wireless Electrocardiogram Signal Acquisition System," in *I2MTC 2016*, 2016.

[9] Y. Li, D. Zhao, W. Serdijn, and others, "A sub-microwatt asynchronous level-crossing adc for biomedical applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 7, no. 2, pp. 149–157, 2013.

[10] J. W. Mark and T. D. Todd, "A nonuniform sampling approach to data compression," *IEEE Transactions on Communications*, vol. 29, no. 1, pp. 24–32, 1981.

[11] K. Kozmin, J. Johansson, and J. Delsing, "Level-crossing ADC performance evaluation toward ultrasound application," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 8, pp. 1708–1719, 2009.

[12] P. G. Steg, S. K. James, D. Atar, L. P. Badano, C. B. Lundqvist, M. A. Borger, C. Di Mario, K. Dickstein, G. Ducrocq, F. Fernandez-Aviles, and others, "ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation," *European heart journal*, vol. 33, no. 20, pp. 2569–2619, 2012.

[13] J. Chen, S. Itoh, and T. Hashimoto, "Wavelet transform based ECG data compression with desired reconstruction signal quality," in *Information Theory and Statistics, 1994. Proceedings., 1994 IEEE-IMS Workshop on*, 1994, p. 84.

[14] I. Daubechies, *Ten lectures on wavelets*. SIAM, 1992.

[15] J.-M. Lina and M. Mayrand, "Complex daubechies wavelets," *Applied and Computational Harmonic Analysis*, vol. 2, no. 3, pp. 219–229, 1995.

[16] U. Lepik, "Numerical solution of differential equations using Haar wavelets," *Mathematics and Computers in Simulation*, vol. 68, no. 2, pp. 127–143, 2005.

[17] J. Shen and G. Strang, "Asymptotics of daubechies filters, scaling functions, and wavelets," *Applied and Computational Harmonic Analysis*, vol. 5, no. 3, pp. 312–331, 1998.

[18] A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Communications on pure and applied mathematics*, vol. 45, no. 5, pp. 485–560, 1992.

[19] 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.

[20] P. E. McSharry, G. D. Clifford, L. Tarassenko, L. Smith, and others, "A dynamical model for generating synthetic electrocardiogram signals,"*IEEE Transactions on Biomedical Engineering*, vol. 50, no. 3, pp. 289–294, 2003.