# ECG SIGNAL DENOISING USING NOISE INVALIDATION

*N.Nikvand*

Dept. of Electrical and Computer Engineering
University of Waterloo
nnikvand@uwaterloo.ca

*H.Bagherzadeh Rafsanjani, M.A Khalilzadeh*

Dept. of Biomedical Engineering
Islamic Azad University- Mashhad Branch
hamid.bagherzadeh@gmail.com,
makhalilzadeh@mshdiau.ac.ir

## ABSTRACT

This paper presents an ECG denoising scheme based on the previously introduced noise invalidation method [1]. The method provides an adaptive data dependent threshold which can be used to denoise ECG signals in an orthogonal basis of choice, by discarding coefficients that are smaller than the proposed threshold. Performance of the noise invalidation method is evaluated using mean square error by denoising a set of ECG signals corrupted by various types of noise, which usually affect ECG signal quality in practice.

***Index Terms—*** ECG, Adaptive Thresholding, Wavelet Transform, Denoising

## 1. INTRODUCTION

The ECG signal is a non-stationary signal which originates in electrical activity of heart and plays a vital role in diagnosis and analysis of heart disease. The signal is usually corrupted by interference coming from electrode movement, power line frequency, muscle noise and other biological signals such as Electromyogram and respiration [2]. To help improve accuracy of the prognosis made by medical experts, it is crucial that these corruptions are eliminated as much as possible at recording stage, using electronic noise reduction techniques and filters, and also at a later stage using signal processing methods.

Several time and frequency domain filtering methods such as Wiener filtering and Kalman filtering, have been proposed to remove these additive corruptions after recording the signal. A review of time and frequency domain filtering methods is found in [3, 4]. However, due to the non-stationarity of the ECG signal, frequency domain filters could cause distortion in a transient interval of the signal and important clinical information maybe lost [5]. To conserve features of signal during transient intervals, many different multi resolution schemes have been proposed, among which wavelet transform is the most popular [2, 6, 7, 5]. The proposed data denoising method in [1] is based on Donoho and Johnstone's wavelet shrinkage which attributes coefficients that are smaller than a certain threshold to noise [8]. Estimated noiseless signal is then reconstructed in time domain using the remaining coefficients. The problem of denoising is then formulated into finding the optimum threshold for the corrupted signal.

## 2. PROBLEM STATEMENT

If a vector of ECG data, $y^N = [y_1, y_2, \cdots, y_N]$ is observed, we are interested to estimate the noiseless data $\bar{y}^N = [\bar{y}_1, \bar{y}_2, \cdots, \bar{y}_N]$ which has been corrupted by a vector of additive noise $w^N = [w_1, w_2, \cdots, w_N]$. We select an orthonormal basis such as $S = [s_1, s_2, \cdots, s_N]$ [1]:

$$< s_i, s_j >= \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{1}$$

Where $s_i$ and $s_j$ are elements of the orthonormal basis. The observed data is then projected into the orthonormal basis of choice:

$$\theta_i =< y^N, s_i > \tag{2}$$

Since the noise is assumed to be additive, it can be shown that:

$$\theta_i = \bar{\theta}_i + \omega_i \tag{3}$$

where $\bar{\theta}_i$ is projection of the noiseless data onto $s_i$ and $\omega_i$ is the respective projection of additive noise. If the noise is further assumed to be an independent and identically distributed gaussian random process, the statistical properties of noise are preserved under orthogonal transforms and we have[2]:

$$E(\Omega_i) = 0, var(\Omega_i) = \sigma^2 \qquad 1 \leq i \leq N \tag{4}$$

In this case, the standard deviation of noise could be estimated using the MAD approach, $\hat{\sigma} = MAD/0.6745$, where MAD is the median of absolute value of normalized fine scale wavelet coefficients of the observed signal and $\hat{\sigma}$ is the estimated standard deviation [9]. The observed coefficients are

[1] $< a, b >$ represents inner product of vectors a & b

[2] We use capital letters to represent random variables and small letters to represent samples of random variables

then thresholded, using the hard or soft thresholding scheme. In [1] soft thresholding is shown to be a better choice with the proposed adaptive scheme and thus we continue by soft thresholding the observed coefficients and reconstructing an estimated version of the noiseless signal using the thresholded coefficients.

$$\hat{\theta}_i = \begin{cases} sign(\theta_i)(|\theta_i| - T_s) & \text{if } \theta_i > T_s \\ 0 & \text{Otherwise} \end{cases} \quad (5)$$

$$\hat{y} = \sum_{i=1}^{N} \hat{\theta}_i s_i \quad (6)$$

Mean Square Error (MSE) is used as a metric of quality of the denoised signal if the original noiseless signal is known. Based on Parseval's theorem, MSE may be calculated in time or scale domain.

$$MSE = \|\bar{y}^N - \hat{y}^N\|^2 = \|\bar{\theta}^N - \hat{\theta}^N\|^2 \quad (7)$$

## 3. NOVEL THRESHOLDING TECHNIQUE

After projecting the observed signal onto the orthogonal basis of choice and deriving the coefficients, we sort the coefficients in descending order based on their absolute value. The ordered coefficients are represented by $\theta_{sort}$ and may be decomposed to noisy and noiseless components:

$$\theta_{sort} = \bar{\theta}_{sort} + \omega_{sort} \quad (8)$$

Where $\bar{\theta}_{sort}$ is noiseless component and $\omega_{sort}$ is the noise component of the observed coefficient. A new variable is defined based on the sorted coefficients as follows:

$$\psi_m = \sum_{i=m+1}^{N} \theta_{sort}[i]^2 = \sum_{i=m+1}^{N} [\bar{\theta}_{sort}[i] + \omega_{sort}[i]]^2 \quad (9)$$

$\Psi_m$ represents a summation of the last $N-m$ absolute sorted squared coefficients in descending order. If a threshold is chosen, the last $N-m$ sorted coefficients are assumed to be noise. In this case, with the assumption that the chosen threshold is optimal, we may have:

$$\psi_m^\omega = \sum_{i=m+1}^{N} \omega_{sort}[i]^2 \quad (10)$$

Where $\psi_m^\omega$ is summation of squared absolute sorted discarded coefficients and is a sample of random variable $\Psi_m^\omega$. The following are the expected value and variance of $\Psi_m^\omega$:

$$E(\Psi_m^\omega) = E(\sum_{i=m+1}^{N} \omega_{sort}[i]^2) \quad (11)$$

$$var(\Psi_m^\omega) = E(\sum_{i=m+1}^{N} \omega_{sort}[i]^2 - E(\sum_{i=m+1}^{N} \omega_{sort}[i]^2))^2 \quad (12)$$

A closed form solution for equations 11 and 12 does not exist. However the expectation and variance could be estimated as follows:

$$\widehat{E(\Psi_m^\omega)} = \frac{1}{J} \sum_{j=1}^{J} \sum_{i=m+1}^{N} z_j[i]^2 \quad (13)$$

$$\widehat{var(\Psi_m^\omega)} = \frac{1}{J-1} \sum_{j=1}^{J} (\sum_{i=m+1}^{N} z_j[i]^2 - \widehat{E(\Psi_m^\omega)})^2 \quad (14)$$

Where $Z$ is a random variable identical to the noise random variable $\Omega_{sort}$. If $\omega_{sort}[i]$'s are samples of independent identically distributed (i.i.d) random variables, central limit theorem may be used to approximate bounds on $\psi_m^\omega$ as follows [1]:

$$Pr\{\frac{|\psi_m^\omega - E(\Psi_m^\omega)|}{\lambda\sqrt{var(\Psi_m^\omega)}} \le 1\} \approx erf(\lambda/\sqrt{2}) \quad (15)$$

Where the error function is

$$erf(x) = \frac{1}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (16)$$

If the samples are not independent or independence of noise samples is lost due to the choice of non-orthogonal basis, a normal bound may still be found on this probability as the number of samples increases, since the random variables are finitely dependent [10] [3].

Equation 15 states that the choice of $\lambda$ would provide an approximation for the probability of $\psi_m^\omega$ being bounded by $E(\psi_m^\omega) \pm \lambda\sqrt{var(\psi_m^\omega)}$. For $\lambda = 5$, we have $erf(\lambda/\sqrt{2}) = 0.999999426697$. A complete discussion of why $\lambda$ is chosen to be 5 is found in [1]. Back to equation 9, each $\psi_m$ is calculated from $N-m$ elements smallest coefficients, sorted in descending order. If we assume that all the elements in $\psi_m$ are noisy elements then we have $\psi_m = \psi_m^\omega$. However starting at some $m^*$ we are going to see the noiseless coefficients coming into the scene and we have $\psi_{m^*} \neq \psi_{m^*}^\omega$, at this point ($m^*$), we can safely select a threshold for which all the $m^*+1$ to $N$ are considered as noise and can be discarded. This process is called Invalidation in [1] and is described as follows:

$$\beta[m] = \frac{|\psi_m - \widehat{E(\psi_m^\omega)}|}{\lambda\sqrt{\widehat{var(\psi_m^\omega)}}} \quad (17)$$

Evidently when $\beta$ is less than one in equation 17, with very high probability we have $\psi_m = \psi_m^\omega$. Once $\beta[m]$ becomes larger than one, depending on the confidence probability we have chosen, in this case $erf(\lambda/\sqrt{2})$, we are certain that $\psi_m$ has noiseless coefficients in its structure. The very first $m$

[3] This property may be used later to remove colored noise from data

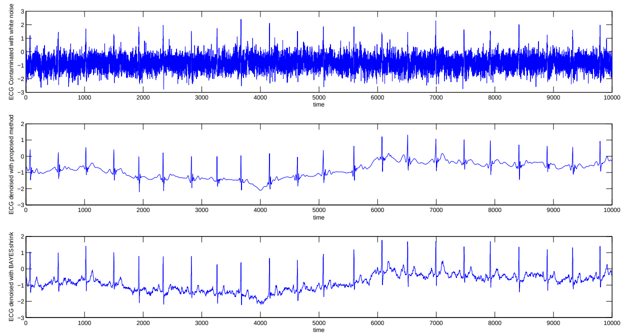

**Fig. 1**. Top: ECG signal contaminated with AWGN, SNR=5, Middle: Denoised signal using the proposed method, Bottom: Denoised signal using BAYESshrink.

that makes $\beta[m]$ larger than one is in fact where the effect of noiseless coefficient is sensed in the $\psi_m$ and the corresponding coefficient may be selected as a proper threshold to separate noise coefficients from actual data.

$$m^* = \arg\min_m (\beta[m] \geq 1) \tag{18}$$

Threshold is thereby selected at $T = |\theta_{sort}[m^*]|$. Obviously the threshold is a function of confidence probability selected through $\lambda$.

## 4. RESULTS

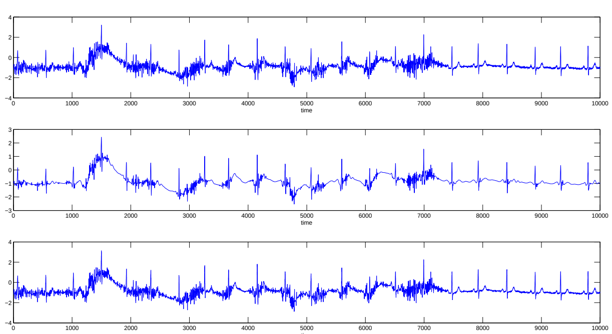

**Fig. 2**. Top: ECG signal contaminated with Muscle artifacts, Middle: Denoised signal using the proposed method, Bottom: Denoised signal using BAYESshrink.

In this section we continue by applying the proposed method to real life ECG data and compare the results to popular thresholding methods, namely VISUshrink, SUREshrink and BAYESshrink. A brief review of the aforementioned methods and further references is available in [1]. Real life data was taken from standard Physiobank [11] data set and

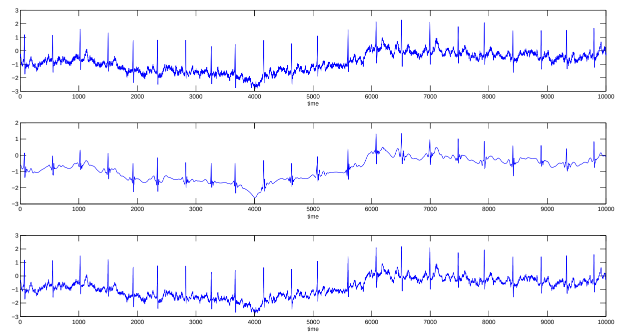

**Fig. 3**. Top: ECG signal contaminated with Colored noise, $\beta = 1.5$, Middle: Denoised signal using the proposed method, Bottom: Denoised signal using BAYESshrink.

five different types of noise was considered. An ECG signal of one minute length was sampled at 360 kHZ frequency (available from Physiobank [11]) and was contaminated with AWGN, muscle noise, colored noise, electrode movement and a mixture of the noises. The noisy signal was then denoised using the proposed method and other Shrink methods. In all cases **'db5'** wavelet was used due to its general resemblance to to the ECG signal and the noisy signal was decomposed up to 8 levels. Figure 1 shows the AWGN contaminated signal, denoised using the proposed method and BAYESshrink. Similarly figures 2 and 3 show contaminated signals denoised by the proposed method and BAYESshrink. Table 1 presents a complete comparison among methods for the aforementioned ECG sample signal contaminated with different noise types. The colored noise that is added to the ECG signal has a $\beta = 1.5$ and the baseline wander weight, electrode movement weight and muscle artifact weight are considered to be one (equal) in the mixture noise. BAYESshrink [12] is widely accepted as a good thresholding technique for signals that have a gaussian or generalized gaussian distribution of coefficients such as mishmash or the ECG signal and it is evident that the proposed method is performing notably well in comparison with this method. In all cases the proposed method shows superior performance compared to the existing methods with notable margin.

**Table 1**. Mean Square Error (MSE) Results

| Case | VISU | SURE | BAYES | PROPOSED |
|---|---|---|---|---|
| AWGN | 0.24 | 0.12 | 0.08 | **0.03** |
| Colored | 0.21 | 0.37 | 0.16 | **0.1** |
| Muscle | 0.38 | 0.21 | 0.18 | **0.12** |
| Baseline | 1.21 | 0.56 | 0.11 | **0.07** |
| Mixture | 1.3 | 0.78 | 0.17 | **0.05** |

## 5. CONCLUSION

A new ECG signal denoising method based on an adaptive thresholding technique was proposed. The method exploits second order statistics of noise to remove additive noise. It was shown that the approach could be generalized to non-gaussian additive noise cases and results of applying the method to the real ECG data contaminated with five different types of practical noise were shown to be promising. While the results of this research could be generalized to other biosignals, it is notable that in general, the thresholding techniques are not the optimum methods for removing artifacts that are caused by natural patient or muscle movements such as electrode movement, baseline wandering caused by respiration, electromyogram noise or a combination of such noises and in such cases using a frequency domain pre-processing step before applying the thresholding may achieve better results. It should also be taken into consideration that mean square error (MSE), the quality measure used in this paper and all the existing literature in this field is not an optimum measure of quality for the denoised signal, and only an expert eye's subjective opinion could rate the true quality of a denoised signal and subsequently a denoising method.

## 6. REFERENCES


[1] S. Beheshti, N. Nikvand, and X.N. Fernando, "Soft thresholding by noise invalidation," in *Communications, 2008 24th Biennial Symposium on*, June 2008, pp. 235–238.

[2] Zhao Zhidong and Pan Min, "ECG denoising by sparse wavelet shrinkage," in *Bioinformatics and Biomedical Engineering, 2007. The 1st International Conference on*, July 2007, pp. 786–789.

[3] O. Sayadi and M.B Shamsollahi, "ECG denoising and compression using a modified extended kalman filter structure," *IEEE Trans. on Biomedical Engineering*, vol. 55, no. 9, pp. 2240–2248, Sept. 2008.

[4] Peter Holland John Yianni Alexander L. Green John F. Stein Tipu Z. Aziz Guohua Lu, John-Stuart Brittainb and Shouyan Wang, "Removing ECG noise from surface emg signals using adaptive filtering," *Neuroscience Letters*, vol. 462, no. 1, pp. 14–19, Sept. 2009.

[5] N. Sivannarayana and D.C. Reddy, "Biorthogonal wavelet transforms for ECG parameters estimation," *Medical Engineering and Physics*, vol. 21, pp. 167–174, 1999.

[6] Mikhled Alfaouri and Khaled Daqrouq, "ECG denoising by sparse wavelet shrinkage," *American Journal of Applied Sciences*, vol. 5(3), pp. 276–281, 2008, [Online `http://www.scipub.org/fulltext/ajas/ajas53276-281.pdf`; accessed 26-June-2010].

[7] O. Sayadi and M.B Shamsollahi, "ECG denoising with adaptive bionic wavelet transform," in *28th Annual International Conf. of IEEE on Eng. in Med. and Biology Society. EMBS '06*, 2006, pp. 6597–6600.

[8] DL Donoho and IM Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.

[9] DL Donoho and IM Johnstone, "adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Assos.*, vol. 90, pp. 1200–1224, 1995.

[10] V. V. Shergin, "central limit theorem for finitely-dependent random variables," *Journal of Mathematical Sciences*, vol. 67, pp. 3244–3248, 1993.

[11] "MIT-BIH," *Physiologic signal archives for biomedical research: The MIT-BIH Noise Stress Test Database*, [Online `http://www.physionet.org/physiobank/database/nstdb/`; accessed 26-June-2010].

[12] Martin Vetterli S.Grace Chang, Bin Yu, "adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing*, vol. 9, pp. 1532–1546, 2000.