Asian Journal of Electrical Sciences
ISSN: 2249 - 6297 Vol. 5 No. 1, 2016, pp.26-35

Noise Estimation and Reduction in Heart Sounds Using Time Frequency Block Thresholding Method

M.Vishwanath Shervegar[1] and Ganesh V Bhat[2]
[1]Assistant Professor, E&C, MIT, Kundapura, Udupi, Karnataka
[2]Principal, CEC, Mangalore, D.K., Karnataka
E-mail: vishwa@hotmail.co.in, ganeshvbhat@yahoo.com

Abstract **- In this paper a novel method of de-noising phonocardiogram by time-frequency Overlapping Group Shrinkage method is described. In this method sigma, the standard deviation of the stationary noise present in a noisy phonocardiogram is found using activity detection. This noise is then canceled by attenuating it in the time frequency domain. The accuracy of noise reduction is measured by SNR. Overlapping Group shrinkage algorithm reduces the effect of noise by attenuating them using hard or soft thresholding. Performance of this method was found to be far better compared to other methods such as Soft Thresholding and Block Thresholding.**
Keywords- **Block thresholding, Activity detection, soft thresholding, overlapping group shrinkage.**

I.INTRODUCTION

Phonocardiogram is a tool used by doctors to look at well-being of any person. Often this phonocardiogram is corrupted by noise from the environment. These noises are usually buzzing and humming sounds from environment, hospital sounds and other artifacts. They hinder the detection of low frequency mild sounds and lead to false detection. So enhancement of phonocardiogram along with noise reduction becomes important. Preliminary literature survey shows that there exist many noise reduction algorithms for phonocardiograms with both merits and demerits. This paper discusses time frequency Overlapping Group Shrinkage algorithm along with soft thresholding and Block Thresholding method. Section 2 discusses heart sound model. Section 3 discusses estimation of heart sound using activity detection. Section 4 discusses the heart sound reduction method using Overlapping Group Shrinkage. Section 5 discusses about the obtained results.

II.HEART SOUND MODEL

Consider a noisy heart sound signal x. It consists of the stationary noise n. The noise is a random noise with an unknown probability density function (pdf) with zero mean [2]. Let the short-time Fourier transform STFT of x be given by (1).

$$X_m^k = \sum_0^{L-1} x[n+mL]e^{-j\frac{2\pi kn}{N}} \quad (1)$$

If we consider that the STFT coefficients of x are a weighted sum of samples of length L of the corresponding random process, then as per central limit theorem, as $L\to\infty$, the STFT coefficients X_m^k asymptotically have Gaussian pdf with zero mean [2]. The pdf of the kth frequency bin X_m^k can be expressed as shown in (2)

$$p(X_m^k) = \frac{1}{\pi\lambda_N(k)} e\{-\frac{|X_m^k|^2}{\lambda_N(k)}\} \quad (2)$$

Thus, the variance of DFT of noise $\lambda_N(k)$ is equivalent to the MMSE estimation of noise power.

The signal x contains heart sound for which $x = s + n$. Activity detection of Phonocardiogram compares the probabilities of presence or absence of heart sound as per the hypothesis stated in (3).

$$\begin{aligned} H_0: X_m^k &= N_m^k,\ \textit{heart sound absent} \\ H_1: X_m^k &= S_m^k + N_m^k, \textit{heart sound present} \end{aligned} \quad (3)$$

Where S_m^k, N_m^k *and* X_m^k are K-dimensional STFT vectors of phonocardiogram (PCG), noise and noisy PCG respectively. Pdf of x, given H_0 is given by (2). Pdf of x, given H_1 is given by (4).

$$p(X_m^k) = \frac{1}{\pi(\lambda_N(k)+\lambda_S(k))} e\{-\frac{|X_m^k|^2}{\lambda_N(k)+\lambda_S(k)}\} \quad (4)$$

Where $\lambda_N(k) = E[|N_m^k|^2]$ and $\lambda_S(k) = E[|S_m^k|^2]$ denote variance of PCG and noise variance respectively.
The likelihood ratio at the kth frequency bin is given by (5) as per the following [4].

$$\Lambda_k = \frac{p(X_k|H_1)}{p(X_k|H_0)} = \frac{1}{1+\xi_k} e^{\left\{\frac{\gamma_k \xi_k}{1+\xi_k}\right\}} \quad (5)$$

$\xi_k = \lambda_S(k)/\lambda_N(k)$ and $\gamma_k = |X_m^k|^2/\lambda_N(k)$ are defined as priori and posteriori snr respectively.

III.ESTIMATION OF NOISE IN PCG

In practice, we do not have an infinite length noise sequence. The most common method of noise estimation

given a finite length noise sequence is periodogram estimation given by (6)[2].
$\hat{\lambda}^m{}_N(k) = |X_m^k|^2$ (6)

where X_m^k is the STFT of noise only signal x in the m[th] frame as defined in (1). We use Bartlett's theorem to reduce the variance $\hat{\lambda}_N^m(k)$ by averaging the M frames.
$\bar{\lambda}_N(k) = \frac{1}{M}\sum_1^{M-1}\hat{\lambda}_N^m(k)$ (6)

This method requires a length LM sequence of noise only observations. $\bar{\lambda}_N(k)$ is an unbiased and consistent estimator of $\lambda_N(k)$: E $[\bar{\lambda}_N(k)] = \lambda_N(k)$ (7)

E $[(\bar{\lambda}_N(k) - \lambda_N(k))^2] = \frac{1}{M}\lambda_N(k)^2$ (8)

However, Eqs. (7) and (8) do not imply that $\bar{\lambda}_N(k)$ predicts any particular instance of $|N_m^k|^2$ with high accuracy: $|N_m^k|^2$is exponentially distributed, so its standard deviation equals its mean.
Considering the PCG presence uncertainty, the MMSE estimate of the noise at the kth frequency bin in the m[th] frame given current noisy observation is mentioned in (2)

$\hat{\lambda}_N^m(k) = E[|N_m^k|^2|X_m^k] = E[|N_m^k|^2|H_0]p(H_0|X_m^k) + E[|N_m^k|^2|H_1]p(H_1|X_m^k)$ (9)

Using Bayes rule, $p(H_0|X_m^k) = \frac{p(X_m^k|H_0)p(H_0)}{p(X_m^k|H_0)p(H_0)+p(X_m^k|H_1)p(H_1)} = \frac{1}{1+\epsilon\Lambda_m^k}$ (10)
Where ϵ = p(H$_1$|H$_0$) and $\Lambda_m^k = p(X_m^k|H_1)|p(X_m^k|H_0)$ is the likelihood ratio of m[th] frame given in (4).
Similarly, $p(H_0|X_m^k) = \frac{\epsilon\Lambda_m^k}{1+\epsilon\Lambda_m^k}$ (11)

If $\beta_k^m = p(H_1|X_m^k)$ then $\hat{\lambda}^m{}_N(k) = \beta_k^m E[|N_m^k|^2|H_1] + (1-\beta_k^m)^2 E[|N_m^k|^2|H_0]$ (12)

Sohn and Sung [7] proposed that, under the hypothesis H$_0$, we can use the current noisy observation, $E[|N_m^k|^2|H_0] = |X_m^k|^2$ (13). Under hypothesis H$_1$, $|X_m^k|^2$ contains PCG as well as noise, and is therefore not an accurate estimate of the noise power.

Assuming that the activity detection of PCG with probability β_k^m has been correctly estimated in all previous frames, the best available estimate of the noise is $E[|N_m^k|^2|H_1] = \hat{\lambda}_N^{m-1}(k)$ (14). From (12), (13) and (14) we have $\hat{\lambda}^m{}_N(k) = \beta_k^m\hat{\lambda}_N^{m-1}(k) + (1-\beta_k^m)^2|X_m^k|^2$. (15)

Sohn and Sung [8] proposed that, if β_k^m is an accurate estimate of the PCG presence probability in each frame, then (15) is an equally accurate estimate of the noise power in the m[th] frame. Under these circumstances, $\hat{\lambda}^m{}_N(k)$ takes into account all information about the underlying noise process that can be extracted from frames up to and including the current frame.

The autoregressive noise estimator $\widehat{\lambda}^m{}_N(k)$ proposed in Eq. (15) is optimal, if and only if the PCG presence probability estimate β_k^m is accurate. Unfortunately, under low SNR conditions $\beta_k^m$is a random variable with high variance. β_k^m is a sigmoid transformation of a random variable $|X_m^k|^2$ given by $\beta_k^m = \frac{e^{\frac{|X_m^k|^2}{a_k\lambda_N(k)}}}{e^{\frac{|X_m^k|^2}{a_k\lambda_N(k)}}+\frac{a_k}{\epsilon}}$ (16), where $a_k = \frac{1+\xi_k}{\xi_k}$. The input threshold to sigmoid function $\theta_k = a_k\lambda_N(k)\log(\frac{a_k}{\epsilon})$ obtained by finding the value of $|X_m^k|^2$ at $\beta_k^m = 0.5$. In noise only frames where $|N_m^k|^2 > \theta_k$ the value of $\beta_k^m = 1$ indicates a false positive even in the absence of PCG. Therefore, autoregressive estimator underestimates noise and overestimates PCG in any given frame. To solve this problem $\beta_k^m$is modelled as binary random variable-a unit step function of $|N_m^k|^2$.Let us define ρ=P ($\beta_k^m \geq 0.5$). The parameter evaluates to $\rho = \int_{a_k\log(\frac{a_k}{\epsilon})}^{\infty} e^{-t}dt = (\frac{a_k}{\epsilon})^{-a_k}$ (17). The expected value of autoregressive estimator approximates to

E $(\hat{\lambda}^m{}_N(k)) = \rho(\hat{\lambda}^{m-1}{}_N(k) + \lambda_N(k)[1-\rho+\rho log\rho]$ (18)

In high noise PCG there is noise propagation error as seen in (16). If the noise process is known to be stationary, and if the first M frames of the signal are known to contain no PCG, then an *a priori* periodogram estimate $\bar{\lambda}_N(k)$ of E $[|N_m^k|^2]$ with known standard error may be computed using Eq. (6). If we assume that intervening frames provide no further information about E $[|N_m^k|^2]$, then

E $[|N_m^k|^2|H_1] = \bar{\lambda}_N(k)$. (19)
$\hat{\lambda}^m{}_N(k) = \beta_k^m\bar{\lambda}_N(k) + (1-\beta_k^m)|X_m^k|^2$ (20)

This method does not propagate error. Instead, a false-positive frame is treated just like any other frame about which we have no certain knowledge of the noise spectrum: the noise estimate is backed off to the *a priori* noise estimator $\bar{\lambda}_N(k)$[8]. The present noise spectrum estimation method can be interpreted as a *posteriori* MMSE estimate of the noise power in the current frame, when the noise process is stationary but with high variance.

IV.TIME FREQUENCY OVERLAPPING GROUP SHRINKAGE ALGORITHM

In recent years, many algorithms based on sparsity have been developed for signal de-noising. These algorithms often utilize nonlinear scalar shrinkage/ thresholding functions of various forms which have been devised so as to obtain sparse representations. Examples of such functions are the hard and soft thresholding functions [8], and the nonnegative garrote [9,10]. Numerous other scalar shrinkage/thresholding functions have been derived as MAP or MMSE estimators using various probability models, e.g. [11,12,13].

For the purpose of de-noising, the regularization parameter λ is chosen analogous to the `three-sigma' rule. The method allows for λ to be selected so as to ensure that the noise variance is reduced to a specified fraction of its original value. This method does not aim to minimize the mean square error or any other measure involving the signal to be estimated, and is thus non-Bayesian. [6]. Let y be the standard deviation of the Gaussian noise in the PCG. Then, y~ N (0,1) and let us define x=soft (y, T). Then the variance of x as a function of threshold T is given by, $\sigma_x^2(T) = E[x^2] = \int_{|y|>T}(|y|-T)^2 p_y(y)\,dy = 2(1+T^2)Q(T) - T\sqrt{\frac{2}{\pi}}e^{\frac{-T^2}{2}}$ (21), where $p_y(y)$ is the standard normal pdf N (0,1) and $Q(T) = \frac{1}{\sqrt{2\pi}}\int_T^\infty e^{\frac{-t^2}{2}}\,dt = 0.5\left(1-\mathrm{erf}\left(\frac{T}{\sqrt{2}}\right)\right)$. [6] Figure 1a shows that standard deviation $\sigma_x(t)$ is a function of threshold T. Soft thresholding uses 3σ rule to attenuate noise. The `3σ rule' states that nearly all values of a Gaussian random variable lie within three standard deviations of the mean (in fact, 99:7%). Since the variance of x is unity here, the 3σ rule suggests setting the threshold to T = 3 which leads to σ_x (3) = 0.020.

The graph in Fig. 1a generalizes the 3σ rule: Given a specified output standard deviation σ_x, the graph shows how to set the threshold T in the soft threshold function so as to achieve it, i.e., so that E [soft (y, T)2] $=\sigma^2_x$ where y ~ N (0, 1). For example, to reduce the noise standard deviation σ to one percent of its value, we solve σ_x(T) = 0.01 for T to obtain T = 3.36σ. This threshold is greater than that suggested by the 3σ rule. OGS provides the alternate solution for this type of problem. In OGS we set the regularization parameter λ as the threshold for PCG detection in the presence of noise. However, for OGS there is no explicit formula such as (21) relating λ to σ_x. Indeed, in the overlapping group case [6], neither is it possible to reduce $E[x^2]$ to a univariate integral as in (21) due to the coupling among the components of y, nor is there an explicit formula for x in terms of y, but only a numerical algorithm. Although no explicit analogue of (21) is available for OGS, the functional relationship can be found numerically. Let y be i.i.d. N (0, 1) and define x as the output of the OGS algorithm: x = ogs(y; λ, K). The output standard deviation σ_x can be found by simulation as a function of λ for a fixed group size. For example, consider applying the OGS algorithm to a two-dimensional array y using a group size of 3X3.

For this group size, σ_x as a function of λ is illustrated in Fig. 1b. The graph is obtained by generating a large two-dimensional array of i.i.d. standard normal random variables, applying the OGS algorithm for a discrete set of λ, and then computing the standard deviation of the result for each λ. Once this graph is numerically obtained, it provides a straight forward way to set λ so as to reduce the noise to a specified level. For example, to reduce the noise standard deviation λ down to one percent of its value, we should use λ=0.43σ in the OGS algorithm according to the graph in Fig. 1b. It can be observed in Fig. 1 that the function σ_x (.) has a sharper `knee' in the case of OGS compared with soft thresholding.

Graphs for numerous group sizes show that in general the larger the group, the sharper is the knee. Note that in practice λ should be chosen large enough to reduce the noise to a sufficiently negligible level, yet no larger so as to avoid unnecessary signal distortion. That is, suitable values of λ are somewhat near the knee. Therefore, due to the sharper knee, the de-noising process is more sensitive to λ for larger group sizes; hence, the choice of λ is more critical. Similarly, it can be observed in Fig. 1 that for OGS, the function $\sigma_x(\lambda)$ follows a linear approximation more closely to the left of the `knee' than it does in the case of soft thresholding.

The preceding sections described how the parameter λ may be chosen so as to reduce additive white Gaussian noise to a desired level. However, in many cases the noise is not white. For example, in the PCG de-noising example where the OGS algorithm is applied directly in the STFT domain. However, the STFT is an over-complete transform; therefore, the noise in the STFT domain will not be white, even if it is white in the original signal domain. In the PCG de-noising the noise is more highly correlated, the values of λ will be somewhat inaccurate.

The penalty function (22) is suitable for stationary noise; however, in many applications, noise is not stationary. For example, in the problem of de-noising PCG corrupted by stationary coloured noise, the Variance of the noise in the STFT domain will vary as a function of frequency. In particular, some noise components may be narrowband and therefore occupy a narrow time-frequency region. The OGS penalty function and algorithm, as described in this paper, do not apply to this problem directly. The penalty function (4) and the process to select λ must be appropriately modified. The OGS algorithm as described above uses the same block size over the entire signal. In some applications, it may be more appropriate that the block size varies. For example, in PCG de-noising, as noted and developed, it is beneficial that the block size in the STFT domain varies as a function of frequency (e.g., for higher temporal resolution at higher frequency). This problem is solved using Block thresholding algorithm.

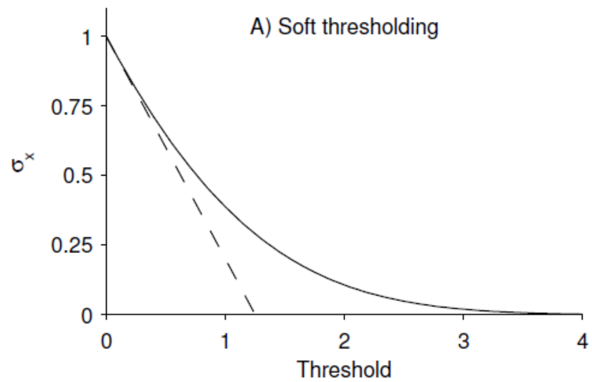

Fig 1a. standard deviation vs threshold

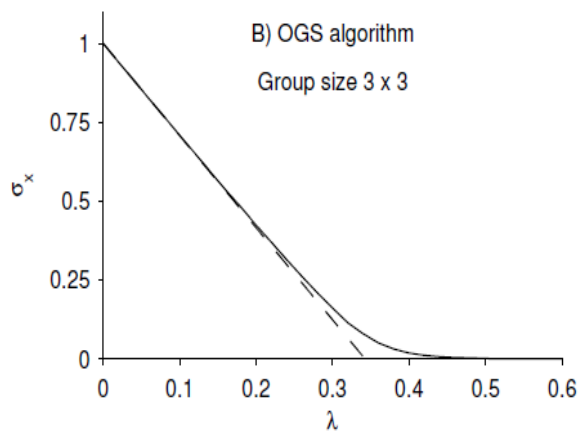

Fig 1b Overlapping group shrinkage (OGS) with group size 3 X 3

V.TIME FREQUENCY BLOCK THRESHOLDING ALGORITHM

A time frequency block estimator regularises the power subtraction by calculating a single attenuation factor. The time frequency plane is divided into I blocks B_i with arbitrary shape. For each B_i a single estimator $\hat{f}$ is calculated with constant attenuation a_i for the noisy signal. The noise characteristics are changed during the passage from time field to time-frequency field. It is still Gaussian (for all frequencies, the noise follows a centred normal law) but σ^2 changes. Consider the discrete Fourier transform of the windowed noise. The Fourier coefficients of the noise is given by $\hat{\eta}_k = \frac{1}{\sqrt{W}} \sum_n w_n \eta_n \exp(\frac{-2i\pi kn}{W})$ (23). Var $(\hat{\eta}_k) = \frac{1}{W} var[\sum_n w(n)\, \eta_n \exp\left(\frac{-2i\pi kn}{W}\right) = 0.375\sigma^2$. (24)

The coefficients matrix is partitioned into macro-blocks and as the signal is real. This matrix has a symmetry between negative and positive frequencies thus it is enough to only treat the negative frequencies. The frequency 0 is treated separately. For the zero frequencies, we treat the points from the beginning to the end eight by eight (blocks 1x8): attenuation coefficient a_i is $a_i = 1 - \frac{1}{\xi_i + 1}$ with $\xi_i = \frac{\bar{Y}_i^2}{\bar{\sigma}_i^2}$. Where $\bar{Y}_i^2$ is the empirical mean on the block i. The real λ is a parameter depending on the block size. The real λ controls the variance term which is due to the noise variation. It is computed with the following expression: $P(\bar{\epsilon}^2 > \lambda\sigma^2) < \delta$. In this expression, δ is a parameter such as, with $\delta = 10^{-3}$, musical noises are barely audible. The blocks inside macro-the block as degree of freedom. Due to discretization effects, λ takes roughly the same values for $W_i = 1$ and $W_i = 2$. So, to compute λ for $W_i = 1$, we are doing the same as if $W_i = 2$. The following matrix gives the computed values of blocks are rectangles. Their sizes are L_i X W_i where L_i and W_i are respectively the length in time and the block width in frequency. The smallest rectangle has the size 1x2, 1 in frequency and 2 in times. With k = 1 (the redundancy factor), $\bar{\epsilon}^2$ is following a χ^2 distribution with the size of λ for different size of blocks (computed thanks to table I):

$$M = \begin{matrix} 1.5 & 1.8 & 2 & 2.5 & 2.5 \\ 1.8 & 2 & 1.5 & 3.5 & 3.5 \\ 2 & 2.5 & 3.5 & 4.7 & 4.7 \end{matrix}$$

TABLE I LAMBDA VS BLOCK SIZE

$B_i^\#$	4	8	16	32	64	128
λ	4.7	3.5	2.5	2.0	1.8	1.5

Even if an upper bound of the risk can be found, then it cannot be computed while the signal f is unknown. That is why we use an estimator of the risk which is found with the SURE theorem. This theorem is used to find the best block shapes into a macro-block by minimising this estimated risk. This is the SURE (Stein Unbiased Risk Estimate) theorem: Let Y be the noisy signal. It's a normal random vector with the identity as covariance matrix and of expectation F, which is the signal searched, without noise. So, Y = F + ε where ε~ N (0, I_p). F is estimated by Y + h (Y), where h is differentiable a.s. h: $R^p \rightarrow R^p$ and $\nabla . h = \sum_{j=1}^{p} \frac{\partial h_j}{\partial y_j}$. Assuming E $[\sum_{j=1}^{p} \left|\frac{\partial h_j}{\partial y_j}\right|] < \infty$, $R = E[\|Y + h(Y) - F\|^2] = p + E[\|h(Y)\|^2 + 2\nabla h(Y)]$. $\hat{R} = p + \|h(Y)\|^2{}_2 + 2\nabla h(Y)$ (25). $\hat{R}$ is an unbiased estimator of the risk R of Y+h(Y). We know that Y_i+h(Y_i)=$a_i Y_i$ is an estimator of F. Therefore, the kth-block risk is:

$$R_k = \sum E[|F(i,j) - a_k F(i,j)|^2] = \sum E[|F(i,j) - Y(i,j) - hY(i,j)|^2] \quad (26)$$

Finally, the blocks are chosen to minimize $\hat{R}_k$. A macroblock is 8 points in time (horizontally) and 16 points in frequency (vertically). The beginning is time 1 and frequency -1. Each macroblock is treated independently. 15 different subdivisions are tested and the best is kept. For a block, the risk can be computed by the following formula

$$\hat{R}_k = \sigma^2 (B_k^\# + \frac{\lambda^2 B_k^\# - 2\lambda(B_k^\# - 2)}{\frac{Y}{\sigma}} 1_{Y \geq \lambda\sigma^2} + B_k^\# \left(\frac{Y}{\sigma} - 2\right) 1_{Y < \lambda\sigma^2}) \quad (27).$$

This formula gives the estimation of the risk of the block i of size $B^\#_i$. It is obtained using the SURE theorem with: p=$B^\#_i$, h(Y_i) = (a_i -1) Y_i. For a given subdivision, the estimation of the risk of the macro-block, is the sum of the risk estimations of each block of the subdivision. All the 15 subdivisions are tested. The one with the minimal risk estimation is chosen. The attenuation coefficients are computed in the same way as for the zero frequency (formula (1)). For the last blocks, which are not full in frequency, all the coefficients of each block are treated together like for the zero frequency. For the last few coefficients that do not make up a block, do hard thresholding. For positive frequencies, conjugate from the negative frequencies.

VI.RESULTS AND DISCUSSION

The experiments presented below have been performed on various types of PCG signals obtained from Peter Bentley's PCG database [1]. The sounds are wav files sampled at 44.1 kHz. They were corrupted by Gaussian noise of different amplitude. For each sound, de-noising with maximum noise removal were applied. The noise power was estimated using activity detection. [2] The database is a mixture of normal and abnormal PCG along with clicks and murmurs. Three methods were used for comparison namely soft thresholding (ST) [6][8], Overlapping group shrinkage (OGS) [6] and Block Thresholding.

Figure 2a shows a noisy heart sound 201108222231 from the database. [1]. Figure 2b illustrates the STFT of the above sound calculated with 50% overlapping blocks of length of 512 samples. A well-known problem arising in many audio enhancement algorithms is that the residual noise is audible as `musical noise' [14] [15]. Musical noise may be attributed to isolated noise peaks in the time-frequency domain that remain after processing. Figure 2c illustrates the STFT obtained by soft thresholding the noisy STFT, with threshold T. T is selected such that as the noise standard deviation reduces down to 0.1% of its value. T = 3.26σ, where σ is the noise standard deviation of the noise in the STFT domain. The noise is sufficiently suppressed and the musical noise is clearly inaudible; however, the signal is distorted due to the relatively high threshold that is used. This is evident from sSNR of 3.07 dB and SNR of 10.91 dB from tables 2 and 3.

Figure 2a shows a noisy heart sound 201108222231 from the database. [1]. Figure 2b illustrates the STFT of the above sound calculated with 50% overlapping blocks of length of 512 samples. A well-known problem arising in many audio enhancement algorithms is that the residual noise is audible as `musical noise' [14] [15]. Musical noise may be attributed to isolated noise peaks in the time-frequency domain that remain after processing. Figure 2c illustrates the STFT obtained by soft thresholding the noisy STFT, with threshold T. T is selected such that as the noise standard deviation reduces down to 0.1% of its value. T = 3.26σ, where σ is the noise standard deviation of the noise in the STFT domain. The noise is sufficiently suppressed and the musical noise is clearly inaudible; however, the signal is distorted due to the relatively high threshold that is used. This is evident from sSNR of 3.07 dB and SNR of 10.91 dB from tables 2 and 3.

Fig. 2c is overly thinned. Methods to reduce musical noise includes over estimating the noise variance, imposing a minimum spectral noise floor [16], and improving the estimation of model parameters [13]. To avoid isolated spurious time-frequency noise spikes (to avoid musical noise), the grouping/clustering behaviour of STFT coefficients of PCG waveforms can be taken into account. To this end, a recent algorithm by Yu et al. [5] for speech/audio enhancement consists of time-frequency block thresholding.

We note that the algorithm [5] is based on non-overlapping blocks. Similar to it [5], the OGS algorithm aims to draw on the grouping behaviour of STFT coefficients so as to improve the overall de-noising result, but it uses a model based on fully overlapping blocks. Figure 2d illustrates the result of block thresholding [3] using the software provided by the authors. It can be seen that block thresholding (BT) produces blocking arte-facts in the spectrogram. Figure 2d illustrates the result of Overlapping Group Shrinkage (OGS) applied to the noisy STFT. 25 iterations of the OGS algorithm were used. Based on listening to HS audio signals de-noised with various group sizes, a group size 8 * 2 (i.e., eight frequency bins * two time bins) was chosen. Other group sizes may be more appropriate for other sampling rates and STFT block lengths. As in the soft thresholding experiment, the parameter λ was selected so as to reduce the noise standard deviation down to 0.1% of its value. Regularization parameter λ was fixed as per λ = 0.32σ. While the sSNR 3.9 dB and SNR 11.9 dB is lower than block thresholding (sSNR 11.81 and SNR 18.63), the artefacts of the OGS de-noised PCG are less audible and musical noise is not audible. This was clearly evident from figure 2e (OGS) and 3d (BT). It was found in [16] that empirical Wiener post-processing (EWP), introduced in [6], improves the result of the block thresholding (BT) algorithm. This post-processing, which is computationally very simple, improves the result of OGS by an even greater degree than for BT, as measured by SNR improvement. The Wiener post-processing raises the SNR for BT from 15.35 dB to 15.75 dB, while it raises the SNR for OGS from 11.9 dB to 16.9 dB while for BT it goes from 18.63 dB to 20.0 dB. Hence, the two methods give almost the same SNR after Wiener post-processing. The substantial SNR improvement in the case of OGS can be explained as follows: the OGS algorithm has the effect of slightly shrinking (attenuating) large coefficients which produces a bias and negatively affects the SNR of the de-noised signal. The Wiener post-processing procedure largely corrects that bias. It has the effect of rescaling (slightly amplifying) the large coefficients appropriately.

TABLE II SNR OF SOUNDS FOR DIFFERENT METHODS

Sounds & SNR (dB)	ST	(OGS-no wiener filtering)	(OGS-wiener filtering)	(BT-no wiener filtering)	(BT-wiener filtering)
201108222231 (-4.24)	10.91	11.9	16.19	18.63	20.0
201108222232 (-9.58)	6.25	7.01	10.43	13.8	15.22
201108222242 (-9.91)	6.15	7.47	12.5	15.39	17.29
201108222257 (-10.33)	7.97	8.21	10.76	13.47	14.75
201108222258 (-3.08)	11.44	11.87	15.96	19.72	21.62
201108222243 (5.12)	17.3	19.78	27.63	18.59	18.68
201108222256 (-14.91)	6.48	6.58	8.49	11.56	12.55

TABLE III SSNR OF SOUNDS FOR DIFFERENT METHODS

Sounds & sSNR (dB)	ST	(OGS-no wiener filtering)	(OGS-wiener filtering)	(BT-no wiener filtering)	(BT-wiener filtering)
201108222231 (-8.81)	3.07	3.5	7.36	11.81	12.93
201108222232 (-9.61)	1.73	2.11	4.6	7.51	8.63
201108222242 (-9.62)	1.93	2.45	5.57	7.9	9.22
201108222257 (-9.66)	1.1	1.16	2.45	4.48	5.33
201108222258 (-8.89)	3.46	3.7	8.27	11.93	13.72
201108222243 (-4.56)	11.73	13.66	26.64	27.31	29.68
201108222256 (-9.87)	0.47	0.48	1.04	1.86	2.27

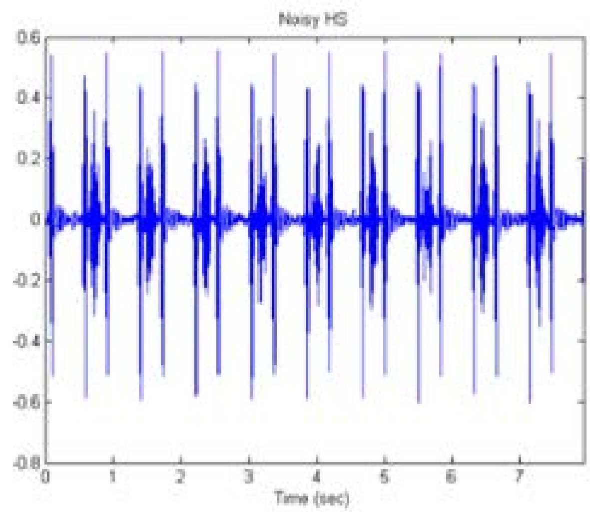

Fig. 2a noisy heart PCG 201108222231

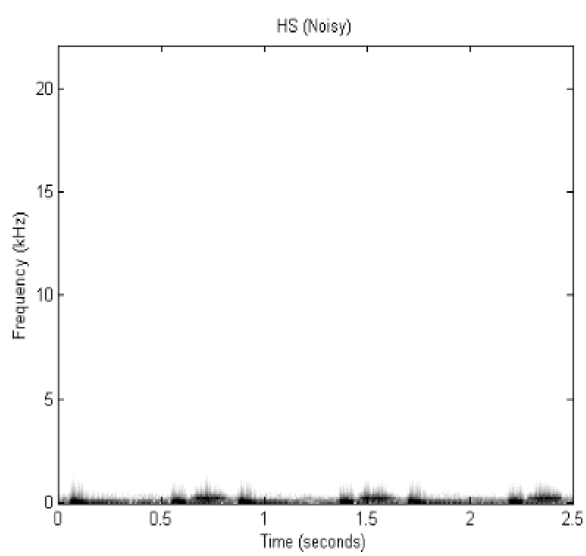

Fig. 2b STFT of noisy signal

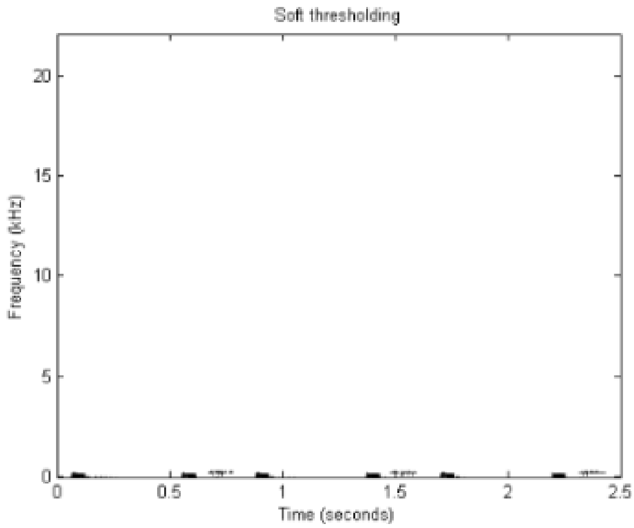

Fig. 2c Soft thresholded PCG

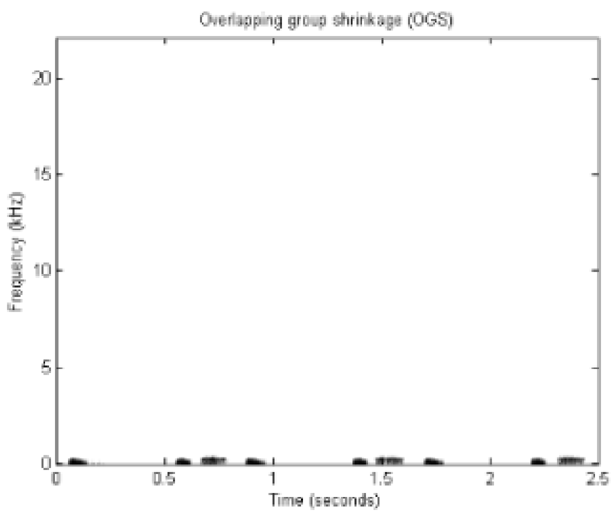

Fig. 2d Overlapping Group Shrinkage PCG

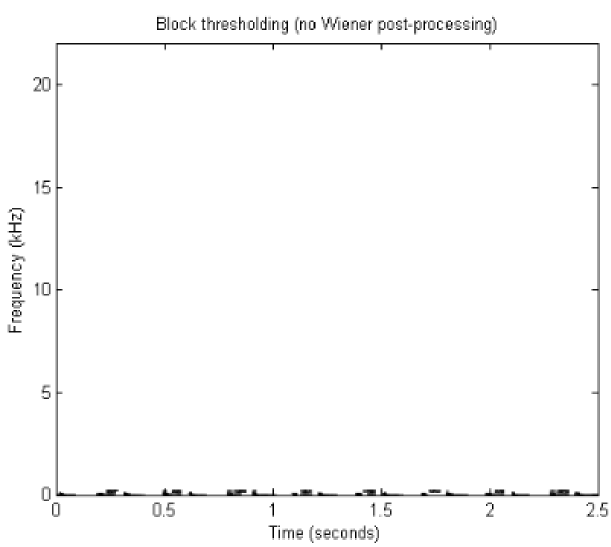

Fig. 2e BT PCG

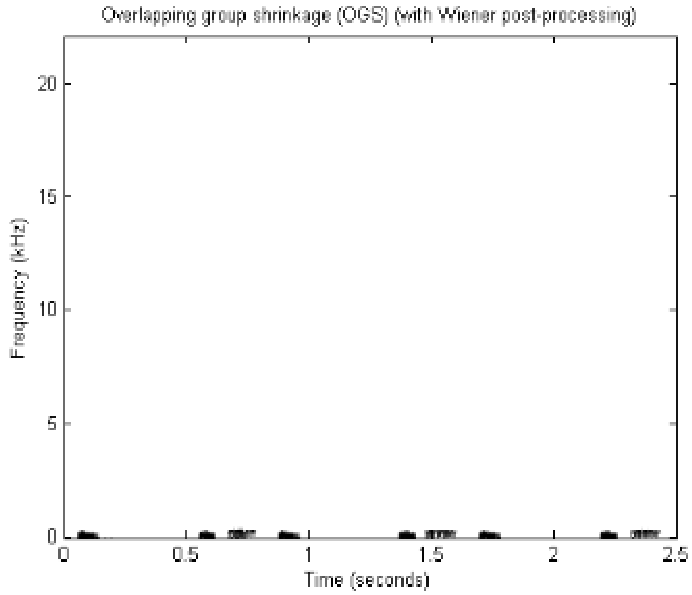

Fig. 2f Wiener Filtered OGS PCG

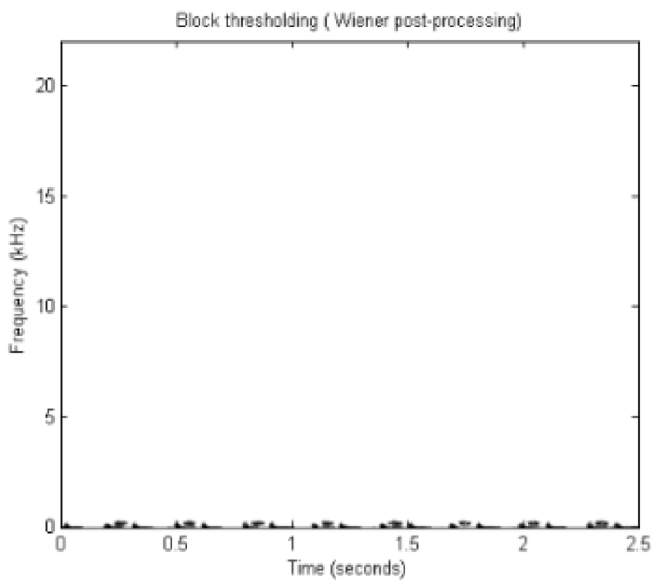

Fig. 2g Wiener Filtered BT PCG

VII.CONCLUSION

From the above work it is clear that both the methods Block thresholding algorithm [3] and OGS algorithm produces de-noised signal with high SNR and sSNR. It is very evident that the so called time frequency structures namely musical noise rarely reoccur in BT and OGS methods. In case of BT algorithm, the SNR and sSNR are quite higher and the sounds are much louder with no artefacts as compared to OGS algorithm. Hence OGS algorithm is the recommended method for de-noising PCG signals especially in hospitals where background noise is a major hindrance in sound acquisition.

REFERENCES

[1] http://www.peterjbentley.com/heartchallenge/
[2] Bowon Lee Hewlett-Packard Laboratories 1501 Page Mill Rd.Palo Alto, CA 94304 bowon.lee@hp.com, Mark Hasegawa-Johnson University of Illinois at Urbana-Champaign Electrical and Computer Engineering 405 N. Mathews Ave,. Urbana, IL61801 jhasegaw@uiuc.edu, MINIMUM MEAN-SQUARED ERROR A POSTERIORI ESTIMATION OF HIGH VARIANCE VEHICULAR NOISE

[3] Guoshen Yu, Stéphane Mallat, *Fellow, IEEE*, and Emmanuel Bacry, Audio Denoising by Time-Frequency Block Thresholding, IEEE Transactions On Signal Processing, Vol. 56, No. 5, May 2008.

[4] Javier Ramírez, José C. Segura, *Senior Member, IEEE*, Carmen Benítez, *Member, IEEE*, Luz García, and Antonio Rubio, *Senior Member, IEEE* Statistical Voice Activity Detection Using a Multiple Observation Likelihood Ratio Test, IEEE Signal Processing Letters, Vol. 12, No. 10, October 2005

[5] Eric Martin, Marie de Masson d'Autume, Christophe Varray, “Audio denoising algorithm with block thresholding”, Published in Image Processing On Line on 2012, July 2 ISSN 2105-1232 ©2012 IPOL.

[6] Po-Yu Chen and Ivan W. Selesnick, Polytechnic Institute of New York University, 6 Metrotech Center, Brooklyn, NY 11201, USA. Email: poyupaulchen@gmail.com, selesi@poly.edu. Tel: +1 718 260-3416. Translation-Invariant Shrinkage/Thresholding of Group Sparse Signals.

[7] J. Sohn and W. Sung, “A voice activity detector employing soft decision based noise spectrum adaptation”, Proc. Int. Conf. Acoust., Speech, and Sig. Process., pp.365-368, 1998.

[8] D. L. Donoho. De-noising by soft-thresholding. IEEE Trans. on Information Theory, 41(3):613{627, May 1995.

[9] M. Figueiredo and R. Nowak. Wavelet-based image estimation: An empirical Bayes approach Using Jerey’s noninformative prior. IEEE Trans. Image Process., 10(9):1322{1331, September 2001.

[10] H. Gao. Wavelet shrinkage denoising using the nonnegative garrote. J. Comput. Graph. Statist., 7:469-488, 1998.

[11] J. M. Fadili and L. Boubchir. Analytical form for a BayesianWavelet estimator of images using the Bessel K form densities. IEEE Trans. Image Process., 14(2):231{240, February 2005.

[12] A. Hyvvarinen. Sparse code shrinkage: Denoising of non-Gaussian data by maximum likelihood estimation. Neural Computation, 11:1739{1768, 1999.

[13] S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans. [14]. S. Boll, “Suppression of acoustic noise in speech using spectral subtraction,” *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. ASSP-27, no. 2, pp. 113–120, Apr. 1979.

[14] S. Boll, “Suppression of acoustic noise in speech using spectral subtraction,” *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. ASSP-27, no. 2, pp. 113–120, Apr. 1979.

[15] M. Berouti, R. Schwartz, and J. Makhoul. Enhancement of speech corrupted by acoustic noise. In Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP), volume 4, pages 208{211, April 1979.

[16] S. Ghael, A. M. Sayeed, and R. G. Baraniuk. Improved wavelet denoising via empirical Wiener filtering.In SPIE Tech. Conf. Wavelet Appl. Signal Proc., San Diego, July 1997.