

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/327429929>

Noise estimation and reduction in heart sounds using time frequency block thresholding method

Article in *International Journal of Biomedical Engineering and Technology* · January 2018

DOI: 10.1504/IJBET.2018.094425

CITATIONS

0

READS

13

2 authors, including:



Vishwanath Shervegar

moodlakatte institute of technology

4 PUBLICATIONS 2 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



heart sound classification [View project](#)

Noise estimation and reduction in heart sounds using time frequency block thresholding method

M. Vishwanath Shervegar*

E&C Department,
MIT, Kundapura,
Udupi, Karnataka, India
Email: vishwa@hotmail.co.in
*Corresponding author

Ganesh V. Bhat

Canara Engineering College,
Mangalore, Karnataka, India
Email: 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 cancelled 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.

Reference to this paper should be made as follows: Shervegar, M.V. and Bhat, G.V. (2018) 'Noise estimation and reduction in heart sounds using time frequency block thresholding method', *Int. J. Biomedical Engineering and Technology*, Vol. 28, No. 1, pp.1–17.

Biographical notes: M. Vishwanath Shervegar works as Assistant Professor in MIT Kundapura, Karnataka, India. He obtained his MTech in Digital Electronics and Advanced Communication from MIT Manipal in 2008. He completed his BE in Electronics and Communication from MIT Manipal in 2003.

Ganesh V. Bhat works as principal in CEC Mangalore, Karnataka, India. He completed his PhD in Statistical Signal Processing from Mangalore University in 2011. He obtained his MTech in Biomedical Engineering from MIT Manipal in 1998. He completed his BE in Electronics and Communication Engineering from MIT Manipal in 1996.

1 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 artefacts. 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 (OGS) 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.

2 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 (Lee and Hasegawa-Johnson, 1996). Let the short-time Fourier transform STFT of x be given by equation (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 $L \rightarrow \infty$, the STFT coefficients X_m^k asymptotically have Gaussian pdf with zero mean (Lee and Hasegawa-Johnson, 1996). The pdf of the k th frequency bin X_m^k can be expressed as shown in equation (2):

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

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 equation (3).

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

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 equation (2). Pdf of x , given H_1 is given by equation (4).

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

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 equation (5) as per the following (Ramírez 2005):

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

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

3 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 equation (6) (Lee and Hasegawa-Johnson, 1996).

$$\hat{\lambda}_N^m(k) = |X_m^k|^2 \quad (6)$$

where X_m^k is the STFT of noise only signal x in the mth frame as defined in equation (1). We use Bartlett's theorem to reduce the variance $\hat{\lambda}_N^m(k)$ by averaging the M frames.

$$\bar{\lambda}_N(k) = \frac{1}{M} \sum_1^{M-1} \hat{\lambda}_N^m(k) \quad (7)$$

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)

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

However, equations (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 k th frequency bin in the m^{th} frame given current noisy observation is mentioned in equation (2).

$$\begin{aligned} \hat{\lambda}_N^m(k) &= E\left[|N_m^k|^2 | X_m^k\right] = E\left[|N_m^k|^2 | H_0\right] p(H_0 | X_m^k) \\ &\quad + E\left[|N_m^k|^2 | H_1\right] p(H_1 | X_m^k) \end{aligned} \quad (9)$$

Using Bayes rule,

$$p(H_0 | X_m^k) = \frac{p(X_m^k | H_0) p(H_0)}{p(X_m^k | H_0) p(H_0) + p(X_m^k | H_1) p(H_1)} = \frac{1}{1 + \epsilon \Lambda_m^k} \quad (10)$$

where $\epsilon = p(H_1 | H_0)$ 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 equation (4). Similarly,

$$p(H_0 | X_m^k) = \frac{\epsilon \Lambda_m^k}{1 + \epsilon \Lambda_m^k} \quad (11)$$

If $\beta_k^m = p(H_1 | X_m^k)$, then

$$\hat{\lambda}_N^m(k) = \beta_k^m E\left[|N_m^k|^2 | H_1\right] + (1 - \beta_k^m)^2 E\left[|N_m^k|^2 | H_0\right] \quad (12)$$

Sohn and Sung (1998) proposed that, under the hypothesis H_0 , we can use the current noisy observation,

$$E\left[|N_m^k|^2 | H_0\right] = |X_m^k|^2 \quad (13)$$

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:

$$E\left[|N_m^k|^2 | H_1\right] = \hat{\lambda}_N^{m-1}(k) \quad (14)$$

From equation (12)–(14), we have:

$$\hat{\lambda}_N^m(k) = \beta_k^m \hat{\lambda}_N^{m-1}(k) + (1 - \beta_k^m)^2 |X_m^k|^2 \quad (15)$$

Sohn and Sung (1998) proposed that, if β_k^m is an accurate estimate of the PCG presence probability in each frame, then equation (15) is an equally accurate estimate of the noise power in the m^{th} frame. Under these circumstances, $\hat{\lambda}_N^m(k)$ 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 $\hat{\lambda}_N^m(k)$ proposed in

equation (15) is optimal, if and only if the PCG presence probability estimate β_k^m is accurate. Unfortunately, under low SNR conditions β_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:

$$\beta_k^m = \frac{e^{\frac{|X_m^k|^2}{a_k \lambda_N(k)}}}{e^{\frac{|X_m^k|^2}{a_k \lambda_N(k)} + \frac{a_k}{\epsilon}}} \quad (16)$$

where $a_k = \frac{1 + \xi_k}{\xi_k}$. The input threshold to sigmoid function $\theta_k = a_k \lambda_N(k) \log\left(\frac{a_k}{\epsilon}\right)$ is 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 β_k^m is modelled as binary random variable-a unit step function of $|N_m^k|^2$. Let us define $\rho = P(\beta_k^m \geq 0.5)$. The parameter evaluates to:

$$\rho = \int_{a_k \log\left(\frac{a_k}{\epsilon}\right)}^{\infty} e^{-t} dt = \left(\frac{a_k}{\epsilon}\right)^{-a_k} \quad (17)$$

The expected value of autoregressive estimator approximates to:

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

In high noise PCG, there is noise propagation error as seen in equation (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 equation (6). If we assume that intervening frames provide no further information about $E[|N_m^k|^2]$, then

$$E[|N_m^k|^2 | H_1] = \bar{\lambda}_N(k) \quad (19)$$

$$\hat{\lambda}_N^m(k) = \beta_k^m \bar{\lambda}_N(k) + (1 - \beta_k^m) |X_m^k|^2 \quad (20)$$

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 $\bar{\lambda}_N(k)$ (Donoho 1995). 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.

4 Time frequency overlapping group shrinkage algorithm

In recent years, many algorithms based on sparsity have been developed for signal de-noising. These algorithms often utilise 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 (Donoho 1995), and the nonnegative garrotte (Figueiredo and Nowak, 2001; Gao, 1998). Numerous other scalar shrinkage/thresholding functions have been derived as MAP or MMSE estimators using various probability models (Fadili and Boubchir, 2005; Hyvvarinen, 1999; Mallat, 1979).

For the purpose of de-noising, the regularisation 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 minimise the mean square error or any other measure involving the signal to be estimated, and is thus non-Bayesian (Chen and Selesnick, 2010). Let y be the standard deviation of the Gaussian noise in the PCG. Then, $y \sim N(0,1)$ and let us define $x = \text{soft}(y, T)$. Then the variance of x as a function of threshold T is given by:

$$\sigma_x^2(T) = E[x^2] = \int_{|y|>T} (|y|-T)^2 p_y(y) dy = 2(1+T^2)Q(T) - T\sqrt{\frac{2}{\pi}}e^{-\frac{T^2}{2}} \quad (21)$$

where $p_y(y)$ is the standard normal pdf $N(0,1)$ and $Q(T) = \frac{1}{\sqrt{2\pi}} \int_T^\infty e^{-\frac{t^2}{2}} dt$
 $= 0.5 \left(1 - \text{erf} \left(\frac{T}{\sqrt{2}} \right) \right)$ (Chen and Selesnick, 2010). Figure 1 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 $\sigma_x(3) = 0.020$. The graph in Figure 1 generalises 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 $E[\text{soft}(y, T)^2] = \sigma_x^2$ where $y \sim N(0, 1)$. For example, to reduce the noise standard deviation σ to one percent of its value, we solve $\sigma_x(T) = 0.01$ for T to obtain $T = 3.36\sigma$. 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 regularisation parameter λ as the threshold for PCG detection in the presence of noise. However, for OGS there is no explicit formula such as equation (21) relating λ to σ_x . Indeed, in the overlapping group case (Chen and Selesnick, 2010), neither is it possible to reduce $E[x^2]$ to a univariate integral as in equation (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 equation (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: $x = \text{ogs}(y; \lambda, K)$. 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 3×3 . For this group size, σ_x as a function of λ is illustrated in Figure 2. 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 $\lambda = 0.43\sigma$ in the OGS algorithm according to the graph in Figure 2. It can be observed in Figure 1 that the function $\sigma_x(\cdot)$ 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 Figure 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. Near the origin, $\sigma_x(\lambda)$ is approximated by

$$\sigma_x(\lambda) = -\sqrt{2} \frac{\Gamma\left(\frac{|\tau|}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{|\tau|}{2}\right)} \lambda, \quad \lambda = 0 \quad (22)$$

where $|\tau|$ is the cardinality of the group (K in 1D, K_1K_2 in 2D) matters. $\tau = \{(j1; j2): 0 \leq j1 \leq K1 \leq 1; 0 \leq j2 \leq K2 \leq 1\}$. This can be explained by noting that for $y \sim N(0, \sigma^2)$, the l2 norm of the group follows a chi-distribution with $|\tau|$ degrees (Chen and Selesnick, 2010) of freedom, the mean of which is the slope in equation (31). For small λ , OGS has roughly the effect of soft thresholding the l2 norm of the groups (Chen and Selesnick, 2010).

Figure 1 Standard deviation vs. threshold

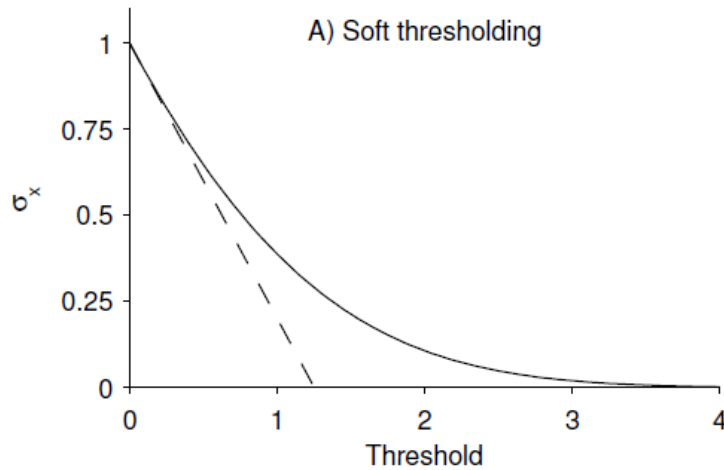
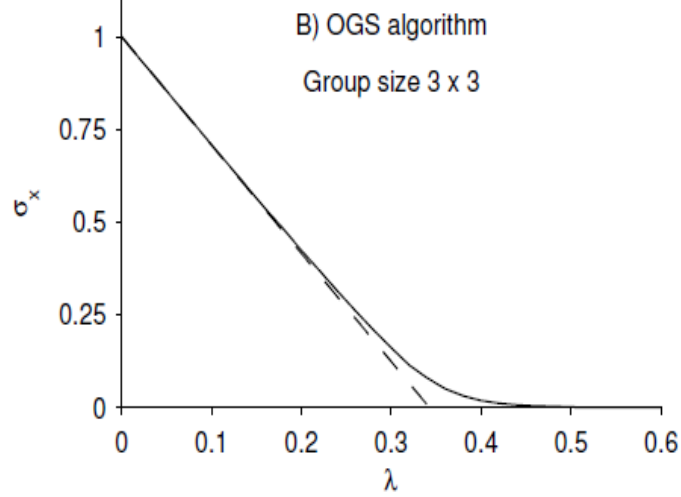


Figure 2 Overlapping group shrinkage (OGS) with group size 3×3 

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 in equation (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 in equation (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.

5 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:

$$\hat{\eta}_k = \frac{1}{\sqrt{W}} \sum_n w_n \eta_n \exp\left(\frac{-2i\pi kn}{W}\right) \quad (23)$$

$$\text{Var}(\hat{\eta}_k) = \frac{1}{W} \text{var} \left[\sum_n w(n) \eta_n \exp\left(\frac{-2i\pi kn}{W}\right) \right] = 0.375\sigma^2 \quad (24)$$

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 1×8): attenuation coefficient a_i is $a_i = 1 - \frac{1}{\xi_i + 1}$ with $\xi_i = \frac{\bar{Y}_i^2}{\sigma_i^2}$. 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-blocks are rectangles. Their sizes are $L_i \times W_i$ where L_i and W_i are, respectively, the length in time and the block width in frequency. The smallest rectangle has the size 1×2 , 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 the block as degree of freedom. Due to discretisation 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 λ for different size of blocks (computed thanks to Table 1):

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

Table 1 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, $Y = F + \varepsilon$, where $\varepsilon \sim N(0, I_p)$. F is estimated by $Y + h(Y)$, where h is differentiable as $h: R^p \rightarrow R^p$ and $\nabla \cdot h = \sum_{j=1}^p \frac{\partial h_j}{\partial y_j}$. Assuming

$$\begin{aligned} E \left[\sum_{j=1}^p \left| \frac{\partial h_j}{\partial y_j} \right| \right] < \infty, R = E[Y + h(Y) - F]^2 \\ \hat{R} = p + E[h(Y)^2 + 2\nabla h(Y)]. \end{aligned} \quad (25)$$

\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 k th-block risk is:

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

Finally, the blocks are chosen to minimise \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

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

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 (equation (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.

6 Results and discussion

The experiments presented below have been performed on various types of PCG signals obtained from Peter Bentley's PCG database (<http://www.peterjbentley.com/heartchallenge/>). 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 (Lee and Hasegawa-Johnson, 1996). 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) (Chen and Selesnick, 2010; Donoho, 1995), Overlapping group

shrinkage (OGS) (Chen and Selesnick, 2010) and Block Thresholding (BT) (Yu and Mallat 2008). Figure 3 shows a noisy heart sound 201108222231 from the database (<http://www.peterjbentley.com/heartchallenge/>). Figure 3 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’ (Boll 1979, Berouti et al. 1979). Musical noise may be attributed to isolated noise peaks in the time-frequency domain that remain after processing. Figure 4 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. $T = 3.26\sigma$, 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. The spectrogram in Figure 4 is overly thinned. Methods to reduce musical noise includes over estimating the noise variance, imposing a minimum spectral noise floor (Ghael 1997), and improving the estimation of model parameters (Mallat, 1979). 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. for speech/audio enhancement consists of time-frequency block thresholding. We note that the algorithm (Martin 2012) is based on non-overlapping blocks. Similar to it (Martin 2012), the OGS algorithm aims to draw on the grouping behaviour of STFT coefficients so as to improve the overall denoising result, but it uses a model based on fully overlapping blocks. Figure 6 illustrates the result of block thresholding (Yu and Mallat 2008) using the software provided by the authors. It can be seen that block thresholding (BT) produces blocking artefacts in the spectrogram. Figure 5 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. Regularisation parameter λ was fixed as per $\lambda = 0.32\sigma$. 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 5 (OGS) and 6 (BT). It was found in Ghael (1997) that empirical Wiener post-processing (EWP), introduced in Chen and Selesnick (2010) 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 18.63dB to 20dB, (Figure 8) while it raises the SNR for OGS from 11.9 dB to 16.9 dB (Figure 7). 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.

Figure 3 Noisy heart sound signal 201108222231 with STFT

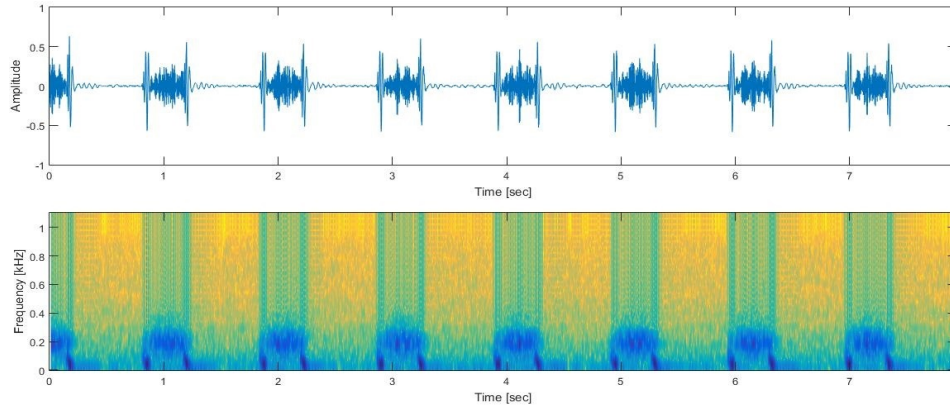


Figure 4 Soft thresholded HS with STFT

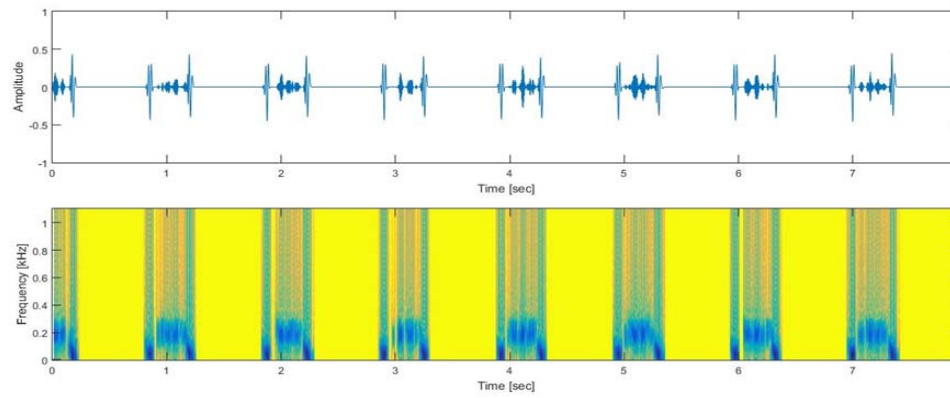


Figure 5 Overlapping Group Shrinkage HS with STFT

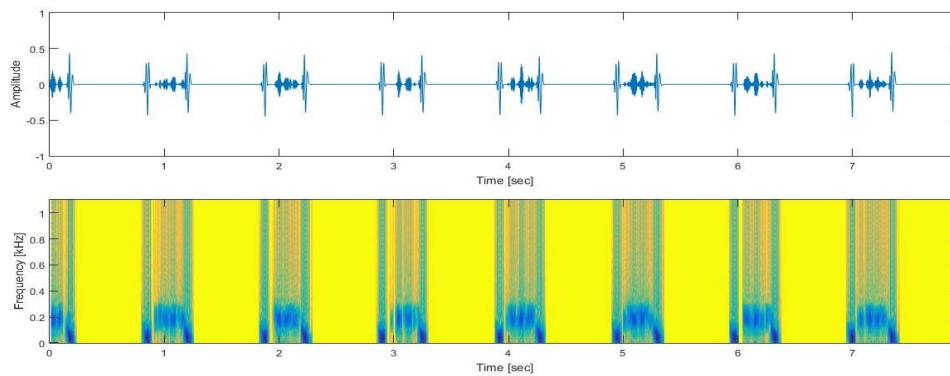


Figure 6 Block thresholded HS with STFT

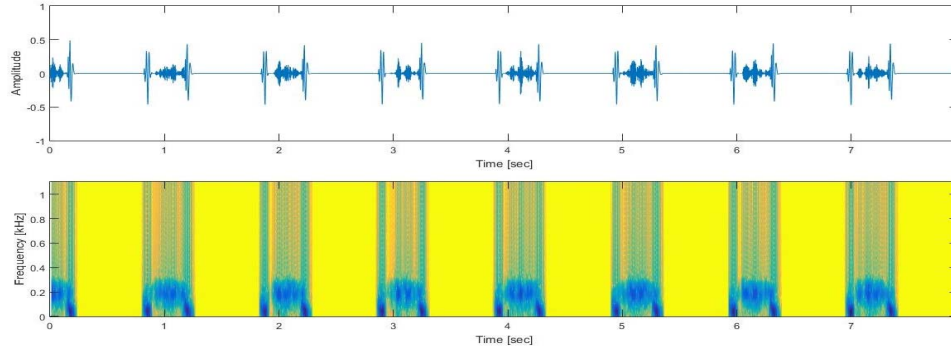


Figure 7 Wiener filtered OGS HS with STFT

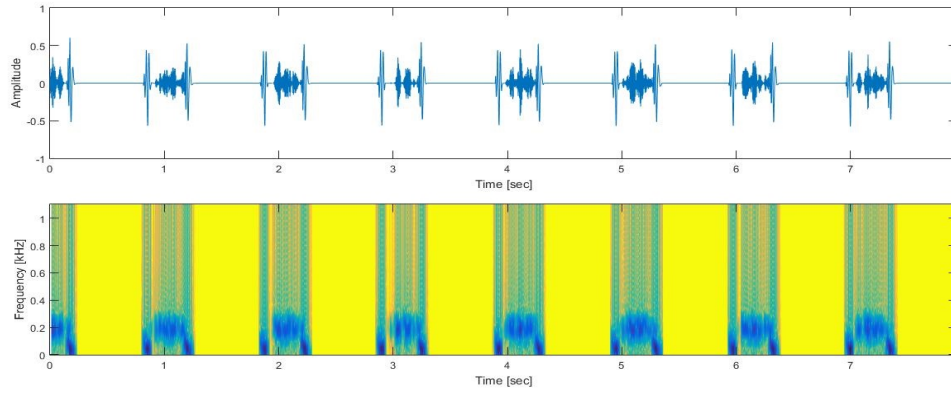


Figure 8 Wiener Filtered BT HS with STFT

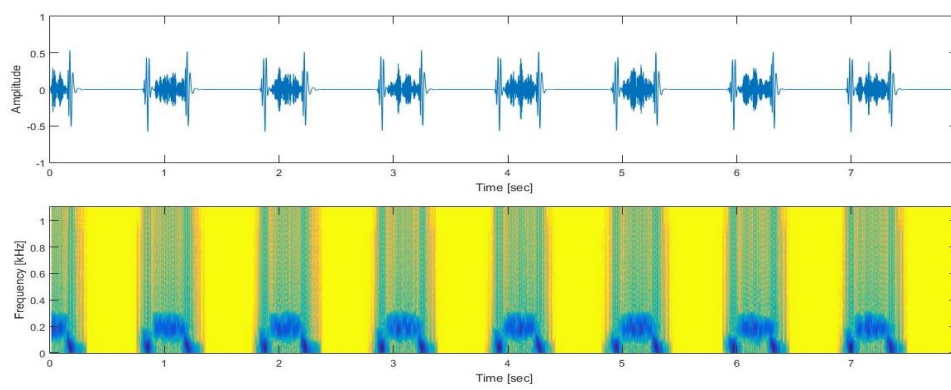


Table 2 SNR of sounds for different methods

<i>Sounds & SNR (dB)</i>	<i>ST</i>	<i>(OGS-no wiener filtering)</i>	<i>(OGS-wiener filtering)</i>	<i>(BT-no wiener filtering)</i>	<i>(BT-wiener filtering)</i>
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 3 SNR of sounds for different methods

<i>Sounds & sSNR (dB)</i>	<i>ST</i>	<i>(OGS-no wiener filtering)</i>	<i>(OGS-wiener filtering)</i>	<i>(BT-no wiener filtering)</i>	<i>(BT-wiener filtering)</i>
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

7 Implications

Phonocardiography is a practice of acquiring, analysing and interpreting the various characteristics and features in heart sound. Often this sound happens to get corrupted with background noise. Today's technology supports various noise filtering techniques that is adopted to remove the redundant noise. With the evolution of digital stethoscope this procedure has got eased out. The various methodologies include adaptive filtering (Mandal et al. 2010), wavelet de-noising, soft thresholding and overlapping group shrinkage method, short term Fourier transform (Djebbari and Reguig 2000). This paper

discusses about time frequency block thresholding method to remove the redundant musical noise that usually appears in other methods. Block thresholding method uses non-diagonal processing methods to remove such noises. At the same time, it improves the SNR of the sounds adequately.

This method has a lot of implication to theory and practice. It will be very helpful for doctors to understand the effect of noise on heart sounds while practising cardiac auscultation (Mandal et al. 2010), foetal phonocardiography (Mithra et al. 2007, Mitra et al. 2010, Chourasia 2012) and bed-side phonocardiography. Cardiac auscultation deals with hearing of fainter and milder as well as high pitched heart sounds. Doctors can now clearly hear murmurs, clicks, snaps and other heart sounds with clarity even amidst the presence of noise. In foetal phonocardiography it will help the doctors to easily distinguish the heart sound of the mother and the baby in the womb. This method will help the doctors to perform real-time monitoring of patients present in the ICU. With the advancement of technology, the method of phonocardiography has gone wireless (Fariborzi and Moghavvemi 2009) and the procedure of noise filtering has become robust. There are many methods in the fray, block thresholding is one such method. Block thresholding method could also aid teaching of cardiac auscultation in hospitals. It will help train doctors about the usage of various methods involved in auscultation and phonocardiography. De-noising by block thresholding method is a precursor to many signal processing methods (Kaur and Singh 2015) such as segmentation and classification. Preliminary literature survey shows that de-noising has improved segmentation accuracies in case of methods in (Wang et al. 2005, Rendón 2013). Block thresholding will improve their results further. Block thresholding method can also be used in analysis and synthesis (Zhang et al 1998) of phonocardiogram signal.

The present method of Block thresholding method has some limitations. Block thresholding method uses block sizes of rectangular shapes for the purpose of de-noising. This does not remove 100% noises. In future, the rectangular shapes can be replaced with different other shapes so that the method becomes more effective and can contain noise better. This paper includes heart sounds taken from Peter Bentley's database. However, the method can be tested under real time hospital conditions. This will give a better birds' eye view about robustness of the described method. Also by incorporating the method in a digital stethoscope the method can be automated to give instantaneous and accurate results.

8 Conclusion

From the above work it is clear that both the methods Block thresholding algorithm (Yu and Mallat 2008) 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

- Boll, S. (1979) 'Suppression of acoustic noise in speech using spectral subtraction', *IEEE Trans. Acoustics, Speech, Signal Process.*, Vol. ASSP-27, No. 2, pp.113–120.
- Berouti, M., Schwartz, R. and Makhoul, J. (1979) 'Enhancement of speech corrupted by acoustic noise', *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP)*, Vol. 4, 208–211.
- Chen, P-Y. and Selesnick, I.W. (2010) 'Translation-invariant shrinkage/thresholding of group sparse signals',
- Chourasia, V.S., Tiwari, A.K. and Gangopadhyay, R. (2012) 'Adaptive neuro-fuzzy inference system for antepartum antenatal care using phonocardiography', *International Journal of Biomedical Engineering and Technology*, Vol. 8, No. 4. DOI: <http://dx.doi.org/10.1504/IJBET.2012.046962>
- Djebbari, A. and Reguig, F.B. (2000) 'Short-time Fourier transform analysis of the phonocardiogram signal', *ICECS 2000 – The 7th IEEE International Conference on Electronics, Circuits and Systems*, IEEE, Jounieh.
- Donoho, D.L. (1995) 'De-noising by soft-thresholding', *IEEE Trans. on Information Theory*, Vol. 41, No. 3, pp.613–627.
- Fadili, J.M. and Boubchir, L. (2005) 'Analytical form for a Bayesian Wavelet estimator of images using the BesselK form densities', *IEEE Trans. Image Process.*, Vol. 14, No. 2, pp.231–240.
- Fariborzi, H. and Moghavvemi, M. (2009) 'Wireless monitoring of cardiac activity: A critical review', *International Journal of Biomedical Engineering and Technology*, Vol. 2, No. 1, pp.4–28. DOI: <http://dx.doi.org/10.1504/IJBET.2009.021905>
- Figueiredo, M. and Nowak, R. (2001) 'Wavelet-based image estimation: An empirical Bayes approach using Jerrey's noninformative prior', *IEEE Trans. Image Process.*, Vol. 10, No. 9, pp.1322–1331.
- Gao, H. (1998) 'Wavelet shrinkage denoising using the nonnegative garrote', *J. Comput. Graph. Statist.*, Vol. 7, pp.469–488.
- Ghael, S., Sayeed, A.M. and Baraniuk, R.G. (1997) 'Improved wavelet denoising via empirical Wiener filtering', *SPIE Tech. Conf. Wavelet Appl. Signal Proc.*, San Diego.
- Hyvvarinen, A. (1999) 'Sparse code shrinkage: Denoising of non-Gaussian data by maximum likelihood estimation', *Neural Computation*, Vol. 11, pp.1739–1768.
- Kaur, S. and Singh, R.M. (2015) 'Survey of different methodologies used in phonocardiogram signal analysis', *International Journal of Computer Applications*, Vol. 117, No. 9.
- Lee, B. and Hasegawa-Johnson, M. (1996) 'Minimum mean-squared error a posteriori estimation of high variance vehicular noise.
- Mallat, S. (1979) 'A theory for multiresolution signal decomposition: The wavelet representation', *IEEE Trans.*
- Mandal, S., Manjunatha, M. and Chatterjee, J. (2010) 'Phonocardiogram signal analysis using adaptive line enhancer methods on mixed signal processor', *2010 International Conference on Signal Processing and Communications (SPCOM)*, IEEE, Bangalore. DOI: <http://dx.doi.org/10.1109/SPCOM.2010.5560509>
- Martin, E., d'Autume, M. de M. and Varray, C. (2012) 'Audio denoising algorithm with block thresholding', *Image Processing*, ISSN 2105-1232 ©2012IPOL.
- Mithra, A.K., Shukla, A. and Zadgaonkar, A.S. (2007) 'System simulation and comparative analysis of foetal heart sound de-noising techniques for advanced phonocardiography', *Int. J. of Biomedical Engineering and Technology*, Vol. 1, No. 1, pp.73–85. DOI: <http://dx.doi.org/10.1504/IJBET.2007.014138>
- Mittra, A.K., Choudhari, N.K. and Zadgaonkar, A.S. (2010) 'Frequency-domain characterisation of fetal heart sound signals: An innovative approach', *International Journal of Biomedical Engineering and Technology*, Vol. 4, No. 1. DOI: <http://dx.doi.org/10.1504/IJBET.2010.034278>

- Ramírez, J., Segura, J.C., Benítez, C., García, L. and Rubio, A. (2005) 'IEEE statistical voice activity detection using a multiple observation likelihood ratio test', *IEEE Signal Processing Letters*, Vol. 12, No. 10.
- Rendón, S.M., Hoyos, C.C., Travieso-Gonzales, C.M. and Castellanos-Domínguez, G. (2013) 'Phonocardiography signal segmentation for telemedicine environments', *Advances in Computational Intelligence*, pp.124–134, DOI: 10.1007/978-3-642-38682-4_15.
- Sohn, J. and Sung, W. (1998) 'A voice activity detector employing soft decision based noise spectrum adaptation', *Proc. Int. Conf. Acoust., Speech, and Sig. Process.*, pp.365–368.
- Wang, P. Kim, Y. Ling, L.H. and Soh, C.B. (2005) 'First heart sound detection for phonocardiogram segmentation', *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology*, pp.5519–5522.
- Yu, G., Mallat, S. and Bacry, E. (2008) 'Audio denoising by time-frequency block thresholding', *IEEE Transactions on Signal Processing*, Vol. 56, No. 5.
- Zhang, X., Durand, L-G, Senhadji, L., Lee, H. and Coatrieux, J-L. (1998) 'Analysis-synthesis of the phonocardiogram based on the matching pursuit method', *IEEE Transactions on Biomedical Engineering*, Vol. 45, No. 8, pp.962–971.