CrossMark

Signal Denoising Using Optimized Trimmed Thresholding

Maha A. Hassanein[1] · **Magdy Tawfik Hanna**[2] ·
Nabila Philip Attalla Seif[1] · **Menna T. M. M. Elbarawy**[2]

Abstract The problem of recovering a desired signal from a recording corrupted by a background additive white Gaussian noise is considered. The undecimated wavelet transform of the noisy recorded signal is taken, and the resulting detail coefficients are thresholded for the purpose of extracting the desired signal. Simple techniques exist for performing the thresholding operation such as the hard, soft and trimmed thresholding methods. Donoho and Johnstone developed a method for selecting the threshold value at every resolution level by minimizing the Stein's unbiased risk estimator (SURE) function while adopting the simple thresholding rationale. They next contributed a hybrid scheme which either uses the last mentioned threshold or defaults to a universal threshold value if the wavelet coefficients are sparse. In the present paper a hybrid scheme is proposed where the trimmed thresholding rationale rather than the soft thresholding rationale is adopted. An expression is first derived for the SURE function for the case of trimmed thresholding before applying the optimization technique. Moreover, instead of using a fixed value of the trimming parameter alpha, a heuristic approach is followed for choosing an optimal value of this parameter. A comparative simulation study is carried out including both standard test signals and

✉ Maha A. Hassanein
  hassanein.maha@eng.cu.edu.eg

  Magdy Tawfik Hanna
  mth00@fayoum.edu.eg

  Nabila Philip Attalla Seif
  npaseif@cu.edu.eg

  Menna T. M. M. Elbarawy
  mmm08@fayoum.edu.eg

1 Department of Engineering Mathematics and Physics, Cairo University, Giza 12613, Egypt

2 Department of Engineering Mathematics and Physics, Fayoum University, Fayoum 63514, Egypt

Birkhäuser

electrocardiogram signals. The simulation results testify to the merit of the contributed method. They show an improvement in the signal-to-noise ratio of the denoised signals extracted by the proposed scheme over those obtained by the universal threshold with hard thresholding and the hybrid SURE threshold with soft thresholding or any non-wavelet-based technique such as the short-time Fourier transform block thresholding, the spectral subtraction or the phase spectrum compensation.

Keywords Signal denoising · Stein's unbiased risk estimator (SURE) · Trimmed thresholding · Hard thresholding · Soft thresholding · Undecimated wavelet transform (UWT)

1 Introduction

The work of Donoho and Johnstone [10] is the foundation of wavelet-based shrinkage and thresholding methods that have been extensively developed for denoising signals [2,4,6,9,15,27]. The wavelet thresholding methods are used in many applications including denoising images [17], electrocardiogram (ECG) signals [1] and speech signals [19,21]. Despite their success, sometimes the traditional wavelet transforms exhibit visual artifacts. Coifman and Donoho [7] proposed the use of the translation-invariant (TI) wavelet transform—known as undecimated wavelet transform (UWT)—for denoising signals. This helps in suppressing the artifacts and consequently producing better results than the discrete wavelet transform (DWT) as measured by the mean squared error (MSE) and signal-to-noise ratio (SNR) [20,22,25].

Although the hard thresholding rule is the simplest one, it can be unstable, i.e., sensitive to small changes in data, because it is defined as a discontinuous function. On the other hand, the soft thresholding rule contributed by Donoho and Johnstone [11]—despite being the most commonly used one—creates unnecessary bias in cases when the true coefficients are large. In order to achieve a compromise between the hard and soft thresholding rules, Fang and Huang [12] proposed the trimmed thresholding rule with a constant parameter α. The trimmed thresholding rule—defined as a continuous function—does not create excessive bias when the wavelet coefficients are large. Actually, the hard and soft thresholding rules are limiting cases of the trimmed thresholding one corresponding to $\alpha \to \infty$ and $\alpha = 1$, respectively. One can achieve the best denoising result by a careful tuning of the trimming parameter α for the signal at hand.

An important step in applying any thresholding method is the determination of an appropriate value of the threshold. Many approaches have been proposed for calculating the threshold value including the mini-max, universal and Stein's unbiased risk estimator (SURE) [24] thresholds. The SureShrink denoising method—which uses a hybrid of the universal threshold and the SURE based on soft thresholding—has been widely used and shown to perform well [5,11]. A wavelet thresholding denoising method that uses a hybrid of the universal threshold and the SURE threshold based on a trimmed thresholding strategy for the UWT coefficients was proposed in [16].

In the present paper, a detailed rigorous derivation of the hybrid SURE with optimized trimmed thresholding strategy is given. Moreover, an optimal value of the trimmed

thresholding parameter α is computed using a simple heuristic technique. In order to illustrate the effectiveness of the proposed method, it is compared with hard thresholding, the SureShrink and the hybrid SURE with trimmed thresholding using a fixed value of α. It is also compared with non-wavelet-based methods such as the short-time Fourier transform (STFT) block thresholding [26], the spectral subtraction [3] and the phase spectrum compensation [23] methods. The extensive simulation study includes both the standard test signals of Donoho and Johnstone [10] and ECG signals. It shows an improvement in the signal-to-noise ratio (SNR) of the denoised signals extracted by the proposed method.

In Sect. 2 the signal denoising problem is formulated and wavelet-based extraction methods are reviewed. In Sect. 3 the hybrid SURE with soft thresholding (SureShrink) is surveyed. In Sect. 4 a detailed derivation of the hybrid SURE with optimized trimmed thresholding and the corresponding implementation algorithm are given. In Sect. 5 the proposed method is applied to test signals as well as ECG signals all contaminated by background additive white Gaussian noise (AWGN). A comparison of the denoising results for different noise levels is made.

2 Wavelet-Based Denoising

A sampled noisy signal $y = (y_1, y_2, \ldots, y_n)^{\mathrm{T}}$ is generated as

$$y = f + \sigma e, \tag{1}$$

where f is a vector of samples of a clean signal, e is a vector of independent identically distributed (i.i.d) zero mean Gaussian noises with unit variance $N(0, 1)$, and σ is the noise level which may be known or unknown. The UWT [18] removes the downsampling operator from the usual implementation of the DWT and upsamples the filter responses, thereby inserting holes between nonzero filter taps. The UWT of the one-dimensional signal y in (1) with sample length 2^J leads to a set of coefficients $\mathcal{W} = \left\{ \hat{c}_{j0,k}, \hat{d}_{j,k} \right\}$, $k = 0, 1, \ldots, 2^J - 1$, $j = j_0, \ldots, J - 1$, where $\hat{c}_{j0,k}$ and $\hat{d}_{j,k}$ are, respectively, the scaling and wavelet coefficients of the noisy data and j_0 is the coarsest resolution level.

Donoho and Johnstone [10] suggested the extraction of the significant wavelet coefficients by thresholding, that is, shrink the wavelet coefficients toward zero if their absolute value is below a certain threshold level, $\lambda \geq 0$. The simplest shrinkage operation is known as "hard thresholding" [10] which is expressed as:

$$\eta_{\mathrm{H}}\left(\hat{d}_{j,k}, \lambda\right) = \begin{cases} \hat{d}_{j,k}, & \left|\hat{d}_{j,k}\right| \geq \lambda \\ 0, & \left|\hat{d}_{j,k}\right| < \lambda, \end{cases} \tag{2}$$

where $\hat{d}_{j,k}$ and $\eta_{\mathrm{H}}\left(\hat{d}_{j,k}, \lambda\right)$ are the empirical wavelet coefficients before and after hard thresholding, respectively, and λ is the threshold level. This is a keep-or-kill approach. The hard thresholding process, however, causes discontinuity in the amplitude of

the shrunk coefficients. To prevent discontinuity, another approach known as "soft thresholding" was introduced in [7] and can be expressed as:

$$\eta_S\left(\hat{d}_{j,k}, \lambda\right) = \begin{cases} \mathrm{sgn}\left(\hat{d}_{j,k}\right)\left(\left|\hat{d}_{j,k}\right| - \lambda\right), & \left|\hat{d}_{j,k}\right| \geq \lambda \\ 0, & \left|\hat{d}_{j,k}\right| < \lambda, \end{cases} \tag{3}$$

where $\eta_S\left(\hat{d}_{j,k}, \lambda\right)$ are the wavelet coefficients after the soft thresholding process and $\mathrm{sgn}(.)$ is the signum function. This is a shrink-or-kill approach which has the merit that the artifacts caused by discontinuities—which appeared in the case of hard thresholding—are not observed. The disadvantage is that all other coefficients which represent the original signal are also shrunk causing a decrease in the SNR of the denoised signal. A more general thresholding scheme which incorporates the advantages of both soft thresholding and hard thresholding is called trimmed thresholding [13] and is expressed as:

$$\eta_T\left(\hat{d}_{j,k}, \lambda\right) = \begin{cases} \hat{d}_{j,k}\left(\frac{\left|\hat{d}_{j,k}\right|^\alpha - \lambda^\alpha}{\left|\hat{d}_{j,k}\right|^\alpha}\right), & \left|\hat{d}_{j,k}\right| > \lambda \\ 0, & \left|\hat{d}_{j,k}\right| \leq \lambda, \end{cases} \tag{4}$$

where $\hat{d}_{j,k}$ and $\eta_T\left(\hat{d}_{j,k}, \lambda\right)$ are the empirical wavelet coefficients before and after trimmed thresholding, respectively, λ is the threshold value, and α is a real-valued parameter lying in the range $1 \leq \alpha \leq \infty$. The wavelet coefficients $\eta\left(\hat{d}_{j,k}, \lambda\right)$—obtained by applying any thresholding rule given in (2), (3) or (4)—are used for a selective reconstruction of the function f in (1). The estimated function—denoted by $\hat{f}(t)$—is obtained by simply performing the inverse undecimated wavelet transform (IUWT) of $\{\hat{c}_{j0,k}, \eta\left(\hat{d}_{j,k}, \lambda\right)\}$.

3 The SURE Approach with Soft Thresholding

Donoho and Johnstone [11] have developed a wavelet-based denoising method, denoted by SureShrink, which is based on adaptively choosing threshold values λ_j at every resolution level j with soft thresholding. Given a set $\{X_1, \ldots, X_p\}$ of statistically independent random variables normally distributed as $X_i \sim N(\mu_i, 1)$ for $i = 1, \ldots, p$, the estimator of the mean vector of the clean signal μ denoted by $\hat{\mu}$ is defined by

$$\hat{\mu} = X + g(X), \tag{5}$$

where $g = (g_i)_{i=1}^p$ is a function from $\mathcal{R}^p$ into $\mathcal{R}^p$. Stein showed that for a weakly differentiable $g(X)$, one has:

$$E_\mu\left\{||\hat{\mu}(X) - \mu||^2\right\} = p + E_\mu\left\{||g(X)||^2 + 2\nabla.g(X)\right\}, \tag{6a}$$

where $\nabla.g(X)$ is the divergence of vector $g(X)$ defined by $\nabla.g(X) \equiv \sum_{i=1}^{p} \frac{\partial g_i}{\partial X_i}$. One can infer—from the above equation—that the quantity

$$\text{SURE}(\lambda; X) = p + ||g(X)||^2 + 2\nabla.g(X) \qquad (6b)$$

is an unbiased estimate of the risk: $E_\mu \left\{ \| \hat{\mu}(X) - \mu \|^2 \right\} = E_\mu \{ \text{SURE}(\lambda; X) \}$. For the soft threshold estimator $\hat{\mu}_i(X) = \eta_S(X_i, \lambda)$ [11], the function $g_i(X)$ is defined by:

$$g_i(X) = -X_i I(i) - \lambda \, \text{sgn}(X_i) \, I'(i), \qquad (7)$$

where $I(i)$ and $I'(i)$ are, respectively, the set indicator functions of the sets $\mathcal{I} = \{i : |X_i| \leq \lambda\}$ and $\mathcal{I}' = \{i : |X_i| > \lambda\}$ and λ is a scalar threshold. The above equation implies that $\frac{\partial g_i}{\partial X_i} = -I(i)$ and consequently

$$\nabla.g(X) = \sum_{i \in \mathcal{I}} -1 = -\#\mathcal{I}, \qquad (8)$$

where $\#A$ denotes the cardinality of set A. Since (7) implies that $\| g(X) \|^2 = \sum_{i=1}^{p} [\min(|X_i|, \lambda)]^2$, the SURE function of (6b) reduces to:

$$\text{SURE}(\lambda; X) = p - 2\#\mathcal{I} + \sum_{i=1}^{p} [\min(|X_i|, \lambda)]^2. \qquad (9)$$

Consequently, the problem of minimizing the MSE function reduces to that of minimizing $\text{SURE}(\lambda; X)$ in order to obtain the optimal threshold λ. Extending this principle to the entire set of resolution levels, at each level j, an expression for the SURE threshold λ_j^s is given by:

$$\lambda_j^s = \underset{0 \leq \lambda \leq \lambda_j^U}{\arg \min} \text{SURE}\left(\lambda, \hat{\mathbf{d}}_j\right), \quad j = j_0, \ldots, J-1, \qquad (10)$$

where $\lambda_j^U = \hat{\sigma} \sqrt{2 \ln 2^j}$ is the level-wise universal threshold [8], $\hat{\sigma}$ is a robust estimate of the noise level, and $\hat{\mathbf{d}}_j$ is the vector of empirical wavelet coefficients at level j whose elements are $\hat{d}_{jk}$, $k = 0, \ldots, 2^j - 1$. The main advantage of the SURE threshold is that when the underlying function has jump discontinuities, the reconstructed signal has no artifacts at the location near discontinuity.

Donoho and Johnstone considered a hybrid scheme of the SURE threshold and the universal threshold—denoted by SureShrink—in order to avoid a serious drawback in situations of extreme sparsity of the wavelet coefficients as reported in [11]. This hybrid scheme is based on the following heuristic idea: If the set of empirical wavelet coefficients is judged to be sparsely represented, then the hybrid scheme defaults to the universal threshold λ_j^U [8]; otherwise, the SURE criterion is used to select a threshold value.

In mathematical terms, the SureShrink is expressed as [11]:

$$
\lambda_j^{\mathrm{H}} = \begin{cases} \lambda_j^{\mathrm{U}} & \text{if } \left\| \hat{\mathbf{d}}_j \right\|_2^2 \leq \hat{\sigma}^2 2^{j/2} \left(2^{j/2} + j^{3/2} \right) \\ \lambda_j^{\mathrm{S}} & \text{otherwise} \end{cases}, \quad j = j_0, \ldots, J-1. \quad (11)
$$

The inequality in (11) is a good measure of sparsity [2,11]. It is observed that when using the universal thresholds, the losses seem to be very small for sparse vectors compared to the case of using SURE, i.e., the hybrid method adaptively "keeps" or "shrinks" large coefficients and "kills" small ones.

4 The Hybrid Scheme with Trimmed Thresholding

The main advantage of the SureShrink method is that it combines the advantages of the universal threshold and the SURE threshold. Therefore, it is widely used in many applications such as speech and image denoising. In this section a hybrid scheme based on the SURE with trimmed thresholding and the universal threshold—to be referred to as the SureTrimm—is introduced. The trimmed thresholding Eq. (4) can be rewritten as

$$
\eta_{\mathrm{T}} \left(X_j, \lambda \right) = X_j \left(1 - \frac{\lambda^\alpha}{\left| X_j \right|^\alpha} \right) I'(j) \quad , \quad \alpha \geq 1, \quad (12)
$$

where $I'(j)$ is the set indicator function of the set $\mathcal{I}'$ defined by $\mathcal{I}' \equiv \left\{ j : \left| X_j \right| > \lambda \right\}$ and $X_1, \ldots, X_p$ are independent $N(\mu_i, 1)$, $i = 1, \ldots, p$. Letting $\hat{\mu}_j = \eta_{\mathrm{T}}(X_j, \lambda)$ in (5) and solving for $g_j(X)$ one obtains:

$$
g_j(X) = \eta_{\mathrm{T}}\left(X_j, \lambda \right) - X_j = X_j \left(1 - \frac{\lambda^\alpha}{\left| X_j \right|^\alpha} \right) I'(j) - X_j. \quad (13)
$$

Upon defining the set indicator function $I(j)$ of the set $\mathcal{I} \equiv \left\{ j : \left| X_j \right| \leq \lambda \right\}$, the above equation reduces to:

$$
g_j(X) = -X_j I(j) - \operatorname{sgn}(X_j) \cdot \frac{\lambda^\alpha}{\left| X_j \right|^{\alpha-1}} I'(j). \quad (14)
$$

Taking the partial derivative, one gets—as derived in "Appendix A"— the following formula:

$$
\frac{\partial g_j}{\partial X_j} = -I(j) + \lambda^\alpha \left\{ \frac{(\alpha-1)}{\left| X_j \right|^\alpha} - \frac{2\delta(X_j)}{\left| X_j \right|^{(\alpha-1)}} \right\} I'(j), \quad (15a)
$$

where $\delta(\cdots)$ is the Dirac delta function. By ignoring the case of $X_j = 0$, the above equation simplifies to:

$$\frac{\partial g_j}{\partial X_j} = -I(j) + \frac{\lambda^\alpha (\alpha - 1)}{|X_j|^\alpha} I'(j). \tag{15b}$$

It is straightforward to show that

$$\frac{\partial g_j}{\partial X_i} = 0 \quad , i \neq j. \tag{15c}$$

Using (15b), the divergence of g reduces to:

$$\nabla . g(X) \equiv \sum_{i=1}^{p} \frac{\partial g_i}{\partial X_i}$$

$$= \sum_{i \in \mathcal{I}} -1 + (\alpha - 1) \sum_{i \in \mathcal{I}'} \frac{\lambda^\alpha}{|X_i|^\alpha}$$

$$= -\#\mathcal{I} + (\alpha - 1) \lambda^\alpha \sum_{i \in \mathcal{I}'} \frac{1}{|X_i|^\alpha}. \tag{16}$$

Based on (14), the squared norm can be computed as:

$$\| g(X) \|^2 = \sum_{j=1}^{p} |g_j|^2$$

$$= \sum_{j=1}^{p} \left| X_j I(j) + \mathrm{sgn}(X_j) \frac{\lambda^\alpha}{|X_j|^{\alpha-1}} I'(j) \right|$$

$$= \sum_{j \in \mathcal{I}} |X_j|^2 + \sum_{j \in \mathcal{I}'} \left(\frac{\lambda^\alpha}{|X_j|^{\alpha-1}} \right)^2$$

$$= \sum_{j \in \mathcal{I}} |X_j|^2 + \lambda^{2\alpha} \sum_{j \in \mathcal{I}'} \frac{1}{|X_j|^{2(\alpha-1)}}. \tag{17}$$

For $j \in \mathcal{I}$: $|X_j| \leq \lambda$ and $|X_j|^\alpha \leq \lambda^\alpha$ for $\alpha \geq 1$ resulting in $|X_j| \leq \frac{\lambda^\alpha}{|X_j|^{\alpha-1}}$. For $j \in \mathcal{I}'$: $|X_j| > \lambda$ and $|X_j|^\alpha > \lambda^\alpha$ for $\alpha \geq 1$ resulting in $\frac{\lambda^\alpha}{|X_j|^{\alpha-1}} < |X_j|$. Consequently, (17) reduces to:

$$\| g(X) \|^2 = \sum_{i=1}^{p} \left[\min(|X_i|, \frac{\lambda^\alpha}{|X_i|^{\alpha-1}}) \right]^2 . \tag{18}$$

According to Stein, if $g(X)$ is a weakly differentiable function as in the case of trimmed thresholding, the SURE $(\lambda; X)$ function of (6b) upon utilizing (16) and (18) reduces to:

$$
\mathrm{SURE}(\lambda; X) = p - 2\#\mathcal{I} + 2(\alpha - 1)\lambda^\alpha \sum_{i \in \mathcal{I}'} \frac{1}{|X_i|^\alpha}
$$
$$
+ \sum_{i=1}^{p} \left[\min\left(|X_i|, \frac{\lambda^\alpha}{|X_i|^{\alpha-1}} \right) \right]^2 . \tag{19}
$$

Applying the above formula to the entire set of resolution levels, an expression for the trimmed threshold λ_j^{T} is obtained as:

$$
\lambda_j^{\mathrm{T}} = \underset{0 \leq \lambda \leq \lambda_j^{\mathrm{U}}}{\arg\min} \mathrm{SURE}\left(\lambda; \hat{\mathbf{d}}_j \right), \quad j = j_0, \ldots, J - 1. \tag{20}
$$

Consequently, the threshold λ_j^{HT} of the hybrid scheme based on the SURE with trimmed thresholding and the universal threshold (SureTrimm) reduces to:

$$
\lambda_j^{\mathrm{HT}} = \begin{cases} \lambda_j^{\mathrm{U}} & \text{if } \left\| \hat{\mathbf{d}}_j \right\|_2^2 \leq \hat{\sigma}^2 2^{j/2} \left(2^{j/2} + j^{3/2} \right) \\ \lambda_j^{\mathrm{T}} & \text{otherwise.} \end{cases} \tag{21}
$$

The main difference between the hybrid scheme of the SURE with trimmed thresholding and the universal threshold (SureTrimm)— briefly presented in [16]—and the SureShrink algorithm of [11] is the use of Eq. (21) in place of Eq. (11) in [11], whereas the main difference between the proposed algorithm of the present paper and that of [16] is the use of an optimal α in contrast to using an arbitrarily fixed α in [16].

The solution of the optimization problem (20) follows the same approach adopted in [11]. One starts by rewriting (19) as:

$$
\mathrm{SURE}\left(\lambda; \hat{\mathbf{d}}_j \right) = p - 2\#\mathcal{I} + 2(\alpha - 1) \sum_{i \in \mathcal{I}'} \left(\frac{\lambda}{\left| \left(\hat{\mathbf{d}}_j \right)_i \right|} \right)^\alpha
$$
$$
+ \sum_{i=1}^{p} \left| \left(\hat{\mathbf{d}}_j \right)_i \right|^2 \left[\min\left(1, \left(\frac{\lambda}{\left| \left(\hat{\mathbf{d}}_j \right)_i \right|} \right)^\alpha \right) \right]^2 , \tag{22}
$$

where $\left(\hat{\mathbf{d}}_j \right)_i$ is the ith element of $\hat{\mathbf{d}}_j$. Rearrange the absolute values of the p elements of the wavelet coefficients vector $\hat{\mathbf{d}}_j$ in ascending order to form the nonnegative vector $v_j = \begin{bmatrix} y_1 & \ldots & y_p \end{bmatrix}^T$, where $y_1 \leq y_2 \leq \ldots \leq y_p$. Let ζ be the index number in the set $\{1, \ldots, p\}$ which satisfies $y_\zeta \leq \lambda$ and $y_{\zeta+1} > \lambda$. Consequently, $\mathrm{SURE}\left(\lambda; \hat{\mathbf{d}}_j \right)$ will

be renamed SURE $\left(\lambda ; \boldsymbol{v}_j\right)$ and (22) will be rewritten in the following simpler form:

$$\text{SURE}\left(\lambda ; \boldsymbol{v}_j\right) = p - 2\zeta + 2\left(\alpha - 1\right) \sum_{i=\zeta+1}^{p} \left(\frac{\lambda}{y_i}\right)^{\alpha} + \sum_{i=1}^{\zeta} y_i^2 + \sum_{i=\zeta+1}^{p} y_i^2 \left(\frac{\lambda}{y_i}\right)^{2\alpha}. \tag{23}$$

The corresponding trimmed threshold λ_j^{T} of (20) will take the form:

$$\lambda_j^{\text{T}} = \underset{0 \leq \lambda \leq \lambda_j^{\text{U}}}{\arg\min} \text{SURE}\left(\lambda ; \boldsymbol{v}_j\right), \quad j = j_0, \ldots, J - 1. \tag{24}$$

Finally, the threshold λ_j^{HT} of the hybrid scheme based on the SURE with trimmed thresholding and the universal threshold (SureTrimm) will be obtained from (21). Moreover, one can achieve an overall mean squared error (MSE) lower than that achieved when using a fixed value of α by optimizing the parameter α at the coarsest resolution level j_0 and choosing the threshold accordingly. Toward that goal a heuristic approach is adopted here.

In order to specify a value for $\alpha \geq 1$ that achieves a minimum value of SURE $\left(\lambda ; \hat{\mathbf{d}}_j\right)$ of (22) over each decomposition level j, the rearranged simple form of SURE, namely SURE $\left(\lambda ; \boldsymbol{v}_{j_0}\right)$ of (23), is computed at discrete values of α over a suitably chosen interval $[1, \bar{\alpha}]$, where $\bar{\alpha}$ is an arbitrary upper bound for α. Consequently, the minimum value of SURE over the chosen interval provides the optimal $\hat{\alpha}$ with the corresponding optimal threshold $\lambda_{j_0}^{\text{T}}$. As a result, at each level of decomposition j, a minimum value of SURE $\left(\lambda ; \hat{\mathbf{d}}_j\right)$ of (22) is computed using the optimal $\hat{\alpha}$. After finding the optimal threshold λ_j^{HT} using (21), the trimmed thresholding of (12) using the optimal $\hat{\alpha}$ is applied to the UWT coefficients to get $\eta_{\text{T}}\left(\hat{d}_{jk}, \lambda_j^{\text{HT}}\right)$. Finally, the denoised signal is reconstructed using the IUWT with the coefficients $\{\hat{c}_{j_0,k}, \quad \eta_{\text{T}}\left(\hat{d}_{jk}, \lambda_j^{\text{HT}}\right)\}$. The main steps are summarized in the following algorithm:

Algorithm:
Given the noisy signal samples y, find an estimate of the original signal f according to model (1).
Input arguments: y, J, j_0, $\bar{\alpha}$ (an arbitrary upper bound for α).

Procedure:

i. Compute the forward UWT coefficients of the noisy signal y: $w_{j,k} = \left(\hat{c}_{j_0,k}, \hat{d}_{j,k}\right)$, for $k = 0, 1, \ldots, 2^J - 1$, $j = j_0, \ldots, J - 1$.
ii. Compute the minimum value of SURE $\left(\lambda ; \boldsymbol{v}_j\right)$ of (23) for $j = j_0$, and find the corresponding $\lambda_{j_0}^{\text{T}}$ using (24) at discrete values of α in $[1, \bar{\alpha}]$. Select $\hat{\alpha}$ which achieves the smallest value of SURE.
iii. For $j = j_0, \ldots, J - 1$:
    a. Compute λ_j^{T} that minimizes SURE from Eq. (24) using $\hat{\alpha}$ in (23).

   b. Determine λ_j^{HT}, the hybrid threshold of (21).

   c. Apply trimmed thresholding to the wavelet coefficients to get $\eta_{\text{T}}\left(\hat{d}_{jk}, \lambda_j^{\text{HT}}\right)$
     using (12).

iv. Compute the inverse undecimated wavelet transform (IUWT) of the thresholded
   coefficients $\{\hat{c}_{j0,k}, \quad \eta_{\text{T}}\left(\hat{d}_{jk}, \lambda_j^{\text{HT}}\right)\}$ and use them in reconstructing the denoised
   signal.

This simple algorithm achieves a better output SNR than other wavelet-based methods
as will be illustrated in the next section. The whole process is computationally efficient
because the UWT and IUWT of steps (i) and (iv), respectively, can be performed by
a fast algorithm having a computational load of the order of $n \log (n)$, where n is the
signal size.

5 Comparative Study

In order to illustrate the effectiveness of the proposed algorithm, the four standard test
signals of Donoho and Johnstone [10]: "Doppler," "Bumps," "Blocks" and "Heavi-
Sine" as well as two electrocardiogram signals—"ECG100.dat" and "ECG117.dat"
downloaded from Physionet [14]—are selected. Additive white Gaussian noise
(AWGN) is added to these signals such that the signal-to-noise ratio (SNR) is 5 dB.
The standard well-known SNR definition is:

$$\text{SNR} = 10 \log_{10} \left[\frac{\sum_{l=0}^{n-1} f^2 (l)}{\sum_{l=0}^{n-1} \left[\hat{f}(l) - f(l)\right]^2} \right], \tag{25}$$

where $f(l)$ and $\hat{f}(l)$ are the lth clean and noisy signal samples, respectively. A sample
of size $n = 1024$ for the four standard test signals and a sample of frame length $n_{\text{F}} =
512$ for the ECG signals are chosen, and the undecimated wavelet transform based on
the Symlet8 wavelets with five decomposition levels is applied. The noisy signal is
filtered using the proposed hybrid scheme of the SURE with trimmed thresholding and
the universal threshold (SureTrimm) with both fixed alpha ($\alpha = 2$) and ($\alpha = 6$) and
optimized alpha ($\hat{\alpha}$), and the results are denoted by TT ($\alpha = 2$), TT ($\alpha = 6$) and TT*,
respectively. For the sake of comparison, the following methods: hard thresholding
(HT), hybrid SURE with soft thresholding (ST) [11], spectral subtraction (SS) [3],
phase spectrum compensation (PSC) [23] and block thresholding (BT) [26], are also
applied.

   Figures 1, 2, 3 and 4 show the clean signals in part (a) and the noisy signals with
SNR = 5 in part (b). The corresponding denoised signals are shown in Figs. 5, 6, 7
and 8 where the results of applying the HT and ST are shown in parts (a) and (b),
respectively. Parts (c) and (d) show the results of the TT ($\alpha = 2$) and TT ($\alpha = 6$),
respectively. Part (e) in those figures shows the results of applying the TT*, where the
indicated values of α are the optimal ones that achieve the minimal risk estimate. Parts
(f), (g) and (h) in Figs. 5, 6, 7 and 8 show the denoised signal using the non-wavelet
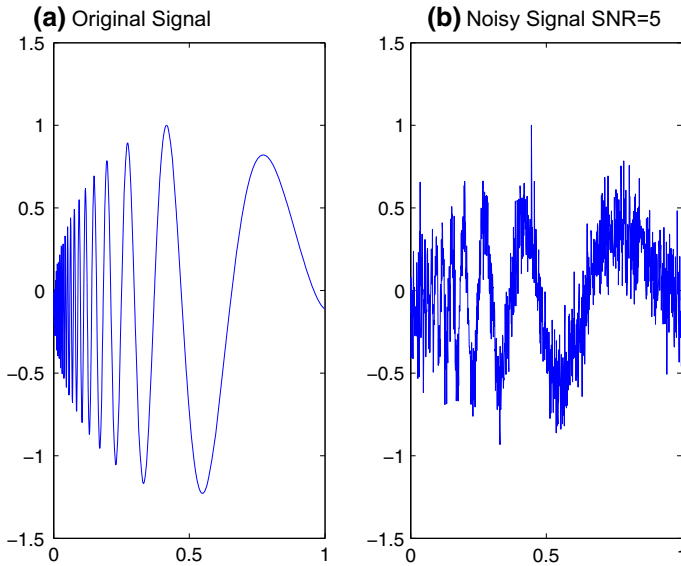
Fig. 1 **a** Doppler signal; **b** noisy signal with SNR = 5 dB
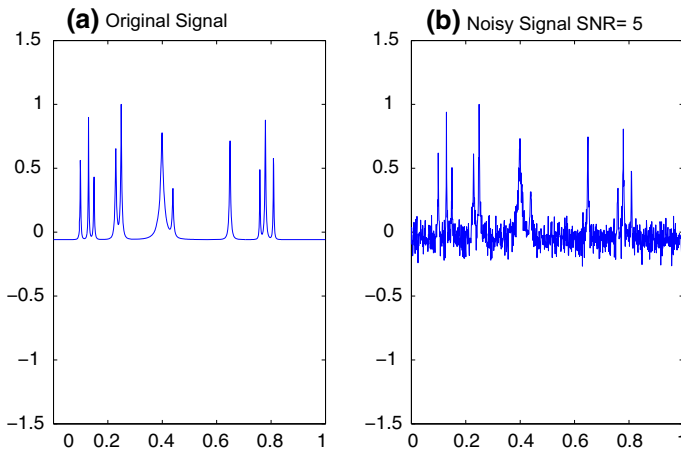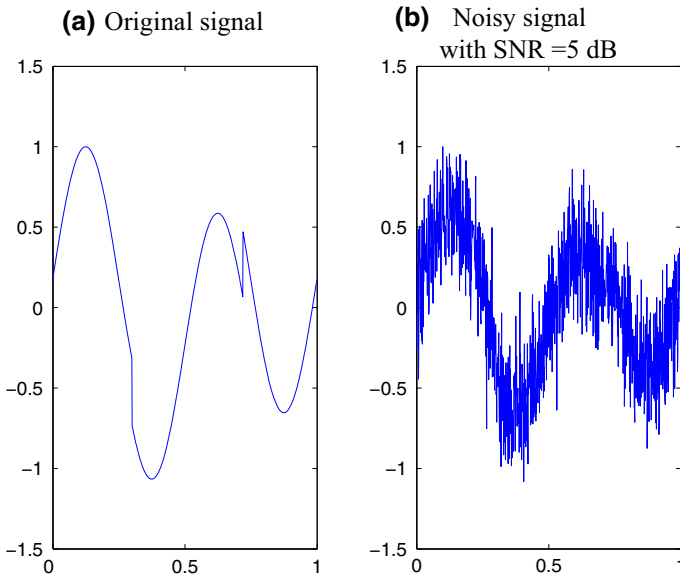

Fig. 2 **a** Bumps signal; **b** noisy signal with SNR = 5 dB

methods: the SS, the PSC and the BT, respectively. An examination of Figs. 5, 6, 7 and 8 reveals the improved performance of the proposed method (TT*) over both wavelet-based and non-wavelet-based denoising methods. The TT* method significantly reduces the noise and preserves the edges of the filtered signal at the same time.

The same methods are applied over a range of input SNRs {− 2, 0, 2, 5, 10, 15} dB, and the SNRs of the denoised signal—computed according to (25)—are listed in Tables 1, 2, 3 and 4 for the four standard test signals. The results testify to the improved
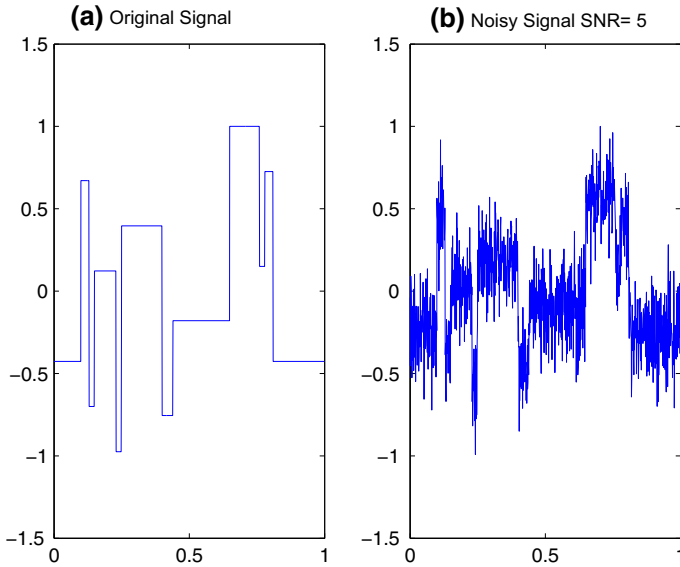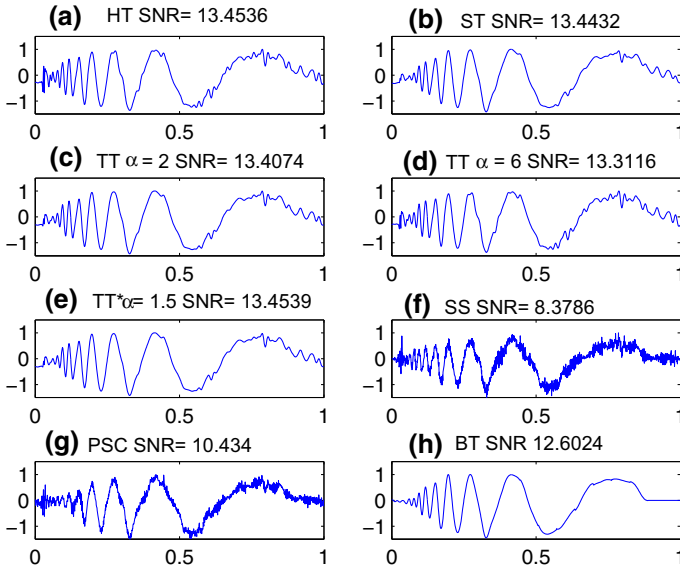
**Fig. 3** **a** Blocks signal; **b** noisy signal with SNR = 5 dB

Fig. 4 **a** HeaviSine signal; **b** noisy signal with SNR = 5 dB

performance of the proposed method (TT*) compared to the other methods for the entire range of input SNR values.

The results of denoising the two electrocardiogram signals—"ECG100.dat" and "ECG117.dat"—are shown in Figs. 9 and 10, respectively. In each of these two fig-
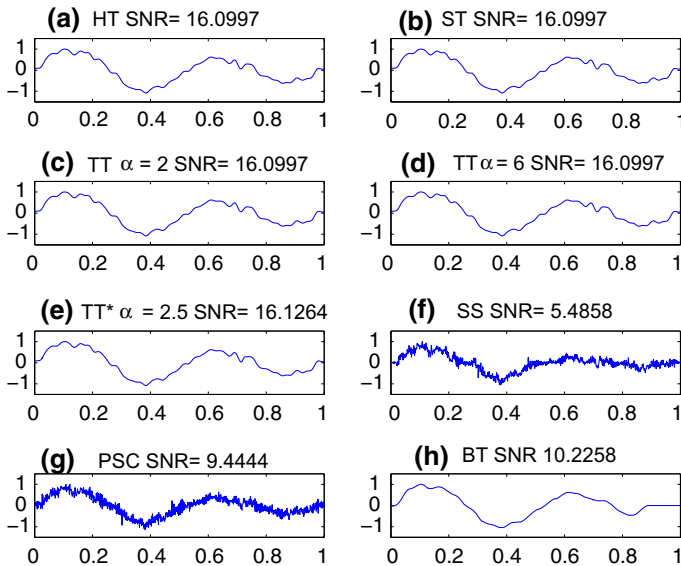
Fig. 5 Doppler signal shown in Fig. 1 denoised using **a** hard thresholding (HT), **b** SureShrink (ST), **c** SureTrimm with fixed alpha TT($\alpha = 2$), **d** TT($\alpha = 6$), **e** SureTrimm with optimized α (TT*), **f** spectral subtraction (SS), **g** phase spectrum compensation (PSC) and **h** block thresholding (BT). (The SNR is in dB.)
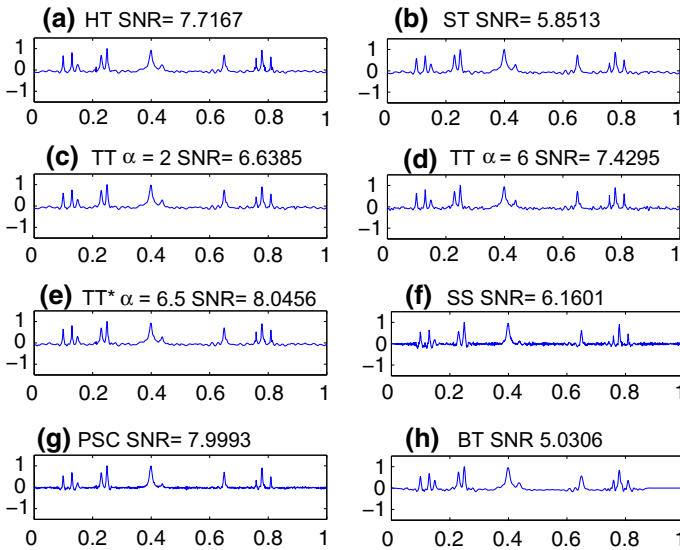
Fig. 6 Bumps signal shown in Fig. 2 denoised using **a** hard thresholding (HT), **b** SureShrink (ST), **c** SureTrimm with fixed alpha TT($\alpha = 2$), **d** TT($\alpha = 6$), **e** SureTrimm with optimized α (TT*), **f** spectral subtraction (SS), **g** phase spectrum compensation (PSC) and **h** block thresholding (BT). (The SNR is in dB.)
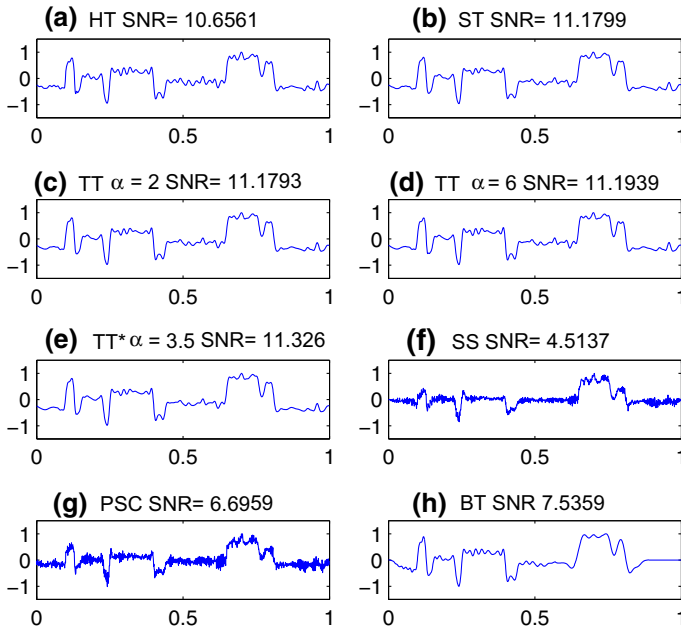
Fig. 7 Blocks signal shown in Fig. 3 denoised using **a** hard thresholding (HT), **b** SureShrink (ST), **c** SureTrimm with fixed alpha TT($\alpha = 2$), **d** TT($\alpha = 6$), **e** SureTrimm with optimized α (TT*), **f** spectral subtraction (SS), **g** phase spectrum compensation (PSC) and **h** block thresholding (BT). (The SNR is in dB.)

Fig. 8 HeaviSine signal shown in Fig. 4 denoised using **a** hard thresholding (HT), **b** SureShrink (ST), **c** SureTrimm with fixed alpha TT($\alpha = 2$), **d** TT($\alpha = 6$), **e** SureTrimm with optimized α (TT*), **f** spectral subtraction (SS), **g** phase spectrum compensation (PSC) and **h** block thresholding (BT). (The SNR is in dB.)

Table 1 Denoising results of Doppler signal

Input SNR	Doppler signal—output SNR							
	HT	ST	TT($\alpha = 2$)	TT($\alpha = 6$)	TT*	SS	PSC	BT
− 2	7.1950	7.1421	7.1493	7.1704	7.2011	5.1569	4.8619	**7.2499**
0	13.5081	13.5173	13.5057	13.5557	**13.6096**	7.8106	6.4830	11.8437
2	10.8353	10.8656	10.8655	10.8651	**10.9207**	6.9849	10.0524	10.4848
5	12.8758	12.8335	12.8339	12.8350	**13.0966**	9.5890	10.0935	12.3967
10	16.9243	17.6272	17.8300	16.0718	**17.8936**	12.4386	11.1976	14.8456
15	19.3361	19.8711	19.9642	19.5798	**20.0214**	15.0164	17.0621	15.8296

Bold values indicate the maximum value in its row
* Signifies an optimized alpha

Table 2 Denoising results of Bumps signal

Input SNR	Bumps signal—output SNR							
	HT	ST	TT($\alpha = 2$)	TT($\alpha = 6$)	TT*	SS	PSC	BT
− 2	5.6270	5.5964	5.6019	5.5019	**5.6464**	3.7986	4.6253	3.1430
0	6.6822	5.5833	6.0509	6.3813	**6.8584**	4.7027	6.2848	5.3020
2	9.2253	8.2171	8.4449	8.9419	**9.3978**	5.0089	4.1556	9.4295
5	8.7934	7.2939	7.8035	8.5060	**8.8306**	6.3884	8.3044	6.2751
10	14.2522	13.3612	14.9509	14.9111	**15.3730**	8.5726	10.9138	10.2752
15	16.6690	15.3977	17.1069	17.2641	**17.7378**	8.8377	15.3686	11.4223

Bold values indicate the maximum value in its row
* Signifies an optimized alpha

Table 3 Denoising results of Blocks signal

Input SNR	Blocks signal—output SNR							
	HT	ST	TT($\alpha = 2$)	TT($\alpha = 6$)	TT*	SS	PSC	BT
− 2	6.2660	6.2806	6.2806	6.2783	**6.3148**	2.6236	3.5540	5.7200
0	8.1212	8.1140	8.1156	8.1198	**8.1261**	3.5085	5.3174	7.1520
2	9.8793	9.8060	9.8115	9.8251	**9.9285**	2.8895	5.0156	6.5788
5	11.3262	11.2199	11.2299	11.2592	**11.3584**	4.4276	7.1618	7.2061
10	13.7021	13.3062	13.4491	13.5817	**13.7791**	4.9138	8.8322	8.1454
15	16.9432	16.8849	17.0871	17.3856	**17.4085**	4.8631	11.3076	8.4811

Bold values indicate the maximum value in its row
* Signifies an optimized alpha

ures, part (a) shows the original ECG signal and part (b) shows the noisy signal with SNR $= 5$ dB. Parts (c)–(h), respectively, show the results obtained by applying the following six methods: HT, ST, TT*, SS, PSC and BT.

Table 4 Denoising results of HeaviSine signal

Input SNR	HeaviSine signal—output SNR							
	HT	ST	TT($\alpha = 2$)	TT($\alpha = 6$)	TT*	SS	PSC	BT
−2	**10.9029**	**10.9029**	**10.9029**	**10.9029**	10.8940	1.6217	4.9864	9.7207
0	11.4281	11.4281	11.4281	11.4281	**11.4282**	2.6899	5.0458	9.6509
2	13.2649	**13.4724**	13.4720	13.4680	13.4273	1.9598	4.6641	10.0382
5	15.7974	**15.9645**	15.9643	15.9643	15.9642	4.4735	7.6952	10.2082
10	19.3538	19.3156	19.3183	19.3273	**19.3598**	3.6143	10.2434	10.4954
15	24.1379	24.0225	24.0345	24.0702	**24.2377**	3.4432	15.0235	10.8011

Bold values indicate the maximum value in its row
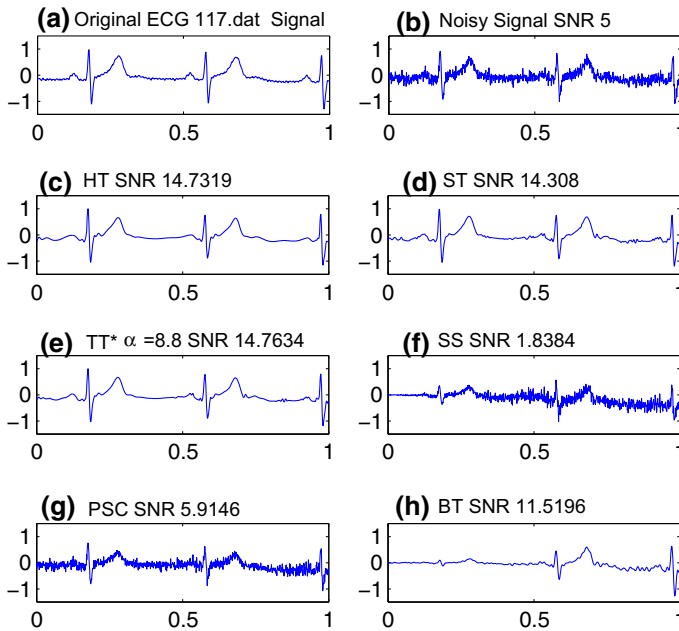* Signifies an optimized alpha
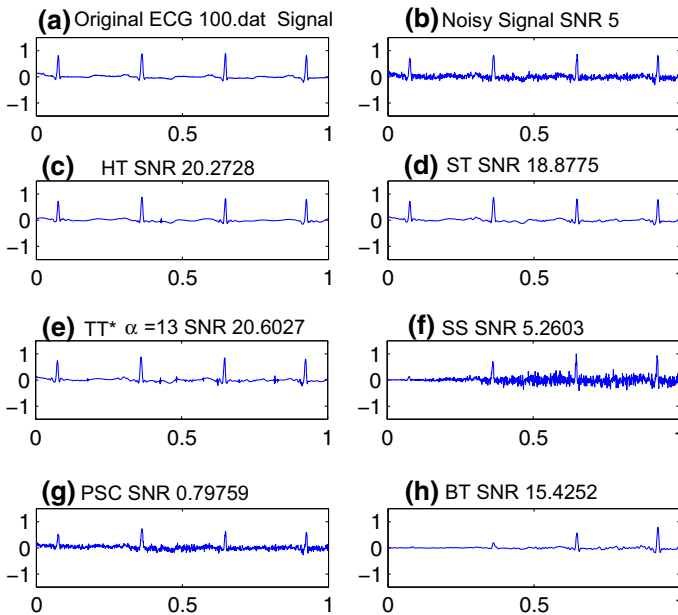
Fig. 9 **a** ECG100 signal, **b** noisy signal, denoised signal using **c** hard thresholding (HT), **d** SureShrink (ST), **e** SureTrimm with optimized α (TT*), **f** spectral subtraction (SS), **g** phase spectrum compensation (PSC) and **h** block thresholding (BT)

6 Conclusion

The problem of extracting a signal from measurements contaminated by additive noise is considered, and a wavelet-based technique is proposed where the undecimated wavelet transform (UWT) detail coefficients are thresholded with the threshold determined according to a hybrid scheme incorporating trimmed thresholding. More

Fig. 10 a ECG117 signal; **b** noisy signal, denoised signal using **c** hard thresholding (HT), **d** SureShrink (ST), **e** SureTrimm with optimized α (TT*), **f** spectral subtraction (SS), **g** phase spectrum compensation (PSC) and **h** block thresholding (BT)

specifically, an expression is derived for the Stein's unbiased risk estimator (SURE) function based on trimmed thresholding and is next minimized in order to determine the threshold value. The philosophy of the hybrid scheme is to either employ the optimal threshold value, or to default to a universal threshold in the case of sparse wavelet detail coefficients. In addition to applying optimization for the determination of the threshold value, a heuristic approach is followed for the selection of the optimal trimmed thresholding parameter alpha. An extensive comparison is made between the proposed method and other wavelet-based methods (the hard thresholding and the hybrid SURE with soft thresholding) and non-wavelet-based methods (spectral subtraction, phase spectrum compensation and block thresholding). The simulation study— including both test signals and ECG signals—demonstrated the merit of the proposed method in terms of the SNR of the denoised signal.

Appendix A

[Derivation of (15a)]

Let

$$h(x) = \operatorname{sgn}(x) \frac{1}{|x|^{(\alpha-1)}}, \tag{A.1}$$

where sgn (x) is the signum function defined by:

$$\text{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0. \end{cases} \tag{A.2}$$

The above equation can be compactly expressed as:

$$\text{sgn}(x) = 2u(x) - 1, \tag{A.3}$$

where $u(x)$ is the unit step function defined by:

$$u(x) = \begin{cases} 1 & x > 0 \\ 0.5 & x = 0 \\ 0 & x < 0. \end{cases} \tag{A.4}$$

The derivative of (A.3) is given by:

$$\frac{d}{dx}\left[\text{sgn}(x)\right] = 2\delta(x), \tag{A.5}$$

where $\delta(x)$ is the Dirac delta function.
Let $\psi(x)$ be the absolute value function defined by:

$$\psi(x) = |x| = \begin{cases} x & x \geq 0 \\ -x & x \leq 0. \end{cases} \tag{A.6}$$

The derivative of the above equation is:

$$\frac{d\psi(x)}{dx} = \begin{cases} 1 & x > 0 \\ -1 & x < 0. \end{cases} \tag{A.7}$$

By comparing (A.2) and (A.7) and ignoring the case of $x = 0$, one gets:

$$\frac{d\psi(x)}{dx} = \text{sgn}(x). \tag{A.8}$$

Based on definition (A.6), one can express (A.1) as:

$$h(x) = \text{sgn}(x) \frac{1}{[\psi(x)]^{(\alpha-1)}}. \tag{A.9}$$

By taking the derivative of the above equation, one obtains:

$$\frac{dh(x)}{dx} = -\text{sgn}(x)(\alpha - 1)[\psi(x)]^{-\alpha}\frac{d\psi}{dx} + \frac{1}{[\psi(x)]^{(\alpha-1)}}\frac{d}{dx}\left[\text{sgn}(x)\right]. \tag{A.10}$$

Upon utilizing (A.5) and (A.8), the above equation reduces to:

$$\frac{\mathrm{d}h\,(x)}{\mathrm{d}x} = \frac{-\,(\alpha - 1)}{[\psi\,(x)]^{\alpha}} + \frac{2}{[\psi\,(x)]^{(\alpha-1)}}\delta\,(x)\,. \qquad (A.11)$$

Based on definition (A.1), one can express (14) as:

$$g_j\,(X) = -X_j\,I\,(j) - \lambda^{\alpha}\mathrm{h}\,(X_j)\,\acute{I}\,(j)\,. \qquad (A.12)$$

By taking the partial derivative of the above equation with respect to X_j and applying (A.11), one gets:

$$\frac{\partial g_j\,(X)}{\partial X_j} = -I\,(j) - \lambda^{\alpha}\left\{\frac{-\,(\alpha - 1)}{\left[\psi\,(X_j)\right]^{\alpha}} + \frac{2}{\left[\psi\,(X_j)\right]^{(\alpha-1)}}\delta\,(X_j)\right\}\acute{I}\,(j)\,. \qquad (A.13)$$

By virtue of (A.6), the above equation reduces to (15a).   □

References

1. M. Alfaouri, K. Daqrouq, ECG signal denoising by wavelet transform thresholding. Am. J. Appl. Sci. **5**(3), 276–281 (2008). doi:10.3844/ajassp.2008.276.281
2. A. Antoniadis, J. Bigot, T. Sapatinas, Wavelet estimators in nonparametric regression: a comparative simulation study. J. Stat. Softw. **6**(6), 1–83 (2001)
3. S. Boll, Suppression of acoustic noise in speech using spectral subtraction. IEEE Trans. Acoust. Speech Signal Process. **27**(2), 113–120 (1979). doi:10.1109/TASSP.1979.1163209
4. T.T. Cai, B.W. Silverman, Incorporating information on neighbouring coefficients into wavelet estimation. Sankhyā Indian J. Stat. Ser. B (1960–2002) **63**(2), 127–148 (2001)
5. S.G. Chang, B. Yu, M. Vetterli, Adaptive wavelet thresholding for image denoising and compression. IEEE Trans. Image Process. **9**(9), 1532–1546 (2000)
6. R. Cohen, *Signal denoising using wavelets, in Department of Electrical Engineering* (Technician, Israel Institute of Technology, Haifa, Israel, 2012)
7. R.R. Coifman, D.L. Donoho, Translation-invariant de-noising, in *Wavelets and Statistics*, vol. 103, Lecture Notes in Statistics, ed. by A. Antoniadis, G. Oppenheim (Springer, New York, 1995), pp. 125–150
8. D.L. Donoho, I.M. Johnstone, *Minimax Estimation Via Wavelet Shrinkage, in Department of Statistics* (Stanford University, Stanford, California, USA, 1992)
9. D.L. Donoho, De-noising by soft-thresholding. IEEE Trans. Inf. Theory **41**(3), 613–627 (1995)
10. D.L. Donoho, J.M. Johnstone, Ideal spatial adaptation by wavelet shrinkage. Biometrika **81**(3), 425–455 (1994). doi:10.1093/biomet/81.3.425
11. D.L. Donoho, J.M. Johnstone, Adapting to unknown smoothness via wavelet shrinkage. J. Am. Stat. Assoc. **90**(432), 1200–1224 (1995)
12. H.-T. Fang, D.-S. Huang, Wavelet de-noising by means of trimmed thresholding, in *The Fifth IEEE World Congress on Intelligent Control and Automation (WCICA 2004) June 2004*, pp. 1621–1624
13. H. Fang, D. Huang, Lidar signal de-noising based on wavelet trimmed thresholding technique. Chin. Opt. Lett. **2**(1), 1–3 (2004)
14. A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, Components of a new research resource for complex physiologic signals. Circulation **101**(23), e215–e220 (2000)
15. D.-F. Guo, W.-H. Zhu, Z.-M. Gao, J.-Q. Zhang, A study of wavelet thresholding denoising, in *5th International Conference on Signal Processing (WCCC-ICSP 2000), Beijing, China August 21–25, 2000*, pp. 329–332

16. M.A. Hassanein, M.T.M.M. Elbarawy, N.P.A. Seif, M.T. Hanna, Trimmed thresholding with sure for denoising signals, in *55th IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), Boise, Idaho, USA August 05–08, (2012)*, pp. 1024–1027

17. P. Hedaoo, S.S. Godbole, Wavelet thresholding approach for image denoising. Int. J. Netw. Secur. Appl. (IJNSA) **3**(4), 16–21 (2011). doi:10.5121/ijnsa.2011.3402

18. M. Holschneider, R. Kronland-Martinet, J. Morlet, P. Tchamitchian, A real-time algorithm for signal analysis with the help of the wavelet transform, in *Wavelets: Time-Frequency Methods and Phase Space, Proceedings of the International Conference, Marseille, France, December 14–18, 1987*, ed. By J.-M. Combes, A. Grossmann, P. Tchamitchian (Springer, Berlin, 1990), pp. 286–297

19. M.T. Johnson, X. Yuan, Y. Ren, Speech signal enhancement through adaptive wavelet thresholding. Speech Commun. **49**(2), 123–133 (2007). doi:10.1016/j.specom.2006.12.002

20. M. Lang, H. Guo, J.E. Odegard, C.S. Burrus, R.O. Wells, Noise reduction using an undecimated discrete wavelet transform. IEEE Signal Process. Lett. **3**(1), 10–12 (1996). doi:10.1109/97.475823

21. F. Qiang, E.A. Wan, Perceptual wavelet adaptive denoising of speech, in *8th European Conference on Speech Communication and Technology (EUROSPEECH), Geneva, Switzerland September 1–4, (2003)*, pp. 1–4. ISCA

22. J.-L. Starck, J. Fadili, F. Murtagh, The undecimated wavelet decomposition and its reconstruction. IEEE Trans. Image Process. **16**(2), 297–309 (2007). doi:10.1109/TIP.2006.887733

23. A.P. Stark, K.K. W'ojcicki, J.G. Lyons, K.K. Paliwal, Noise driven short-time phase spectrum compensation procedure for speech enhancement, in *Interspeech, Brisbane, Australia September 22–26, (2008)*, pp. 549–552

24. M.C. Stein, Estimation of the mean of a multivariate normal distribution. Ann. Stat. **9**(6), 1135–1151 (1981)

25. L. Su, G. Zhao, De-noising of ECG signal using translation-Invariant wavelet de-noising method with improved thresholding, in *The 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE-EMBS 2005), Shanghai, China January 17–18, (2006)*, pp. 5946–5949

26. G. Yu, S. Mallat, E. Bacry, Audio denoising by time-frequency block thresholding. IEEE Trans. Signal Process. **56**(5), 1830–1839 (2008). doi:10.1109/TSP.2007.912893

27. X.-P. Zhang, M.D. Desai, Adaptive denoising based on SURE risk. IEEE Signal Process. Lett. **5**(10), 265–267 (1998). doi:10.1109/97.720560