Computer Science and Information Systems 23(2):687–706 https://doi.org/10.2298/CSIS250828017H

# Joint Heart Sound Denoising Using DTCWT and Adaptive Sparsity-assisted Signal Smoothing Algorithm

Jianqiang Hu[1,2], Dafeng Shen[1], Lin Chen[1], Yu Chen[1], Shigen Shen[3], and Yan Che[2]

[1] School of Computer and Information Engineering, Xiamen University of Technology, Xiamen 361024, P. R. China
hujianqiang@tsinghua.org.cn
shendafeng25@gmail.com
275914270c@gmail.com
631894828@qq.com

[2] Engineering Research Center for Big Data Application in Private Health Medicine of Fujian Universities, Putian University, Putian, Fujian 351100, P. R. China
ptucy07@126.com (corresponding author)

[3] School of Information Engineering, Huzhou University, Huzhou 313000, P. R. China
shigens@zjhu.edu.cn

**Abstract.** Various unwanted and unavoidable noises corrupt Heart Sound Signals (HSSs). It is strongly required to suppress respiratory sound and ambient noise due to significant reduction of clarity and interpretation of HSSs. In this paper, we propose a joint heart sound denoising using Dual-Tree Complex Wavelet Transform (DTCWT) and Adaptive Sparsity-assisted Signal Smoothing (ASASS) algorithm. In this research, the signal is first decomposed by DTCWT to obtain the multi-scale feature representation of the signal. Subsequently, ASASS suppresses pseudo-Gibbs artifacts around signal boundaries of DTCWT while implementing adaptive thresholding strategies to maximize the Signal-to-Noise Ratio (SNR). Experimental validation on the PhysioNet/CinC 2016 database and Open Access Heart Sound Dataset (OAHS Dataset) demonstrates that the proposed method significantly outperforms existing techniques. Under conditions involving Gaussian white noise (GWN) SNR of 0 dB, the proposed method achieves an SNR of 9.01 dB and a Root Mean Square Error (RMSE) of 0.032, outperforming standalone DTCWT and multiple existing models.

**Keywords:** Heart Sound Signals, Denoising, Adaptive Sparsity-Assisted Signal Smoothing, Dual-Tree Complex Wavelet Transform.

## 1. Introduction

Internet of Things (IoT)-based digital stethoscopes are rapidly entering the field of family healthcare monitoring [1]. Heart sounds are collected by various digital stethoscopes, which contain a great number of biomedical signals of cardiac activity. In practice, heart sound is corrupted by unavoidable entities that easily lead to the clinical misinterpretations. Ambient noise, respiratory sound and even signals from various complex environments overlap with the heart sound signals (HSSs), making their quality degraded. In particular, the spectrum of respiratory sound overlaps perfectly with that of heart sound. Besides, electronic interferences and recording artifacts corrupt HSSs to a certain extent.

HSS is a non-linear, non-stationary weak signal with a low frequency and weak amplitude. Obviously, heavy noises bring about the poor the diagnostic interpretations. In order to obtain accurate diagnostic performance, it is necessary to effectively suppress noise contaminants from HSSs. As a result, HSS denoising has become a primary challenge in research and clinical practice.

Over the past few decades, the researchers have developed various effective denoising methods to maximize the preservation of critical components of HSSs while eliminating noises. First, multiple time-domain methods have been employed for noise removal, including conventional filters and auto-correlation techniques. For instance, Butterworth band-pass filters and Finite Impulse Response (FIR) filters can eliminate noises outside the frequency range of HSS [2]. Second, various frequency-domain methods such as Wavelet transform (WT) [3], Fourier transform [4], Empirical Wavelet Transform (EWT) [5], Singular Value Decomposition (SVD) [6] and Dual-Tree Complex Wavelet Transform (DTCWT) [7] have been utilized to for HSSs denoising. Among these, WT-based denoising relies on adaptive thresholding of wavelet coefficients, enabling simultaneous noise suppression and preservation of singularity points in HSSs. Compared to WT, DTCWT offers advantages including time-invariance, superior reconstruction capability, and absence of aliasing effects-exhibiting excellent anti-aliasing performance and near-shift invariance. However, a common challenge across all WT-based denoising methods is the introduction of pseudo-Gibbs artifacts at the singularities. This occurs primarily because the amplitude of local oscillations decreases near signal discontinuities (singularity points). Finally, machine learning and deep learning have been applied to remove noise contaminations from HSSs. Machine learning-based denoising, such as particle Swarm optimization (PSO) [8] and Twin Support Vector Machine [9], demonstrates limited generalization capability when confronted with diverse noise types. While deep learning possesses stronger feature representation capabilities than machine learning, it struggles with handling unbalanced datasets. For instance, loss functions of fully convolutional networks (FCNs) [10], denoising convolutional neural network (DnCNN) [11] and Enhancement Adversarial Convolutional Neural Network (EACNN) [12], remain challenging in practice due to the requirement for maximizing the preservation of subtle details within HSS fluctuations.

In this work, we develop a joint denoising method for HSSs based on DTCWT and Adaptive Sparsity-Assisted Signal Smoothing (ASASS) algorithm, called DTCWT+ASASS. While DTCWT exhibits excellent near-shift invariance, it may still be insufficient for effectively capturing detailed components of HSSs. By integrating ASASS with DTCWT, the proposed approach simultaneously processes low-pass recursive filters and zero-phase non-causal high-pass filters as banded matrices to denoise diverse signals. Notably, the zero-phase characteristic eliminates phase distortion caused by causal Linear Time-Invariant (LTI) filters, thereby preserving the original morphology of HSSs.

The principal contributions of this study are summarized as follows:

1. We propose a joint HSS denoising method using DTCWT and ASASS. This method prevents distortion in denoised HSS while suppressing pseudo-Gibbs artifacts at signal boundaries in DTCWT, consequently enhancing denoising adaptability.
2. We introduce an ASASS algorithm incorporating an adaptive thresholding mechanism. This algorithm smoothens HSSs and suppresses unwanted overlapping noise frequencies to recover noise-free signals. Crucially, its adaptive mechanism dynami-

cally adjusts thresholding strategies based on sub-band energy distribution and directional correlations, achieving optimal balance between noise suppression and feature preservation.

3. To validate the effectiveness of the proposed method, multiple experiments were performed on the PhysioNet/CinC 2016 Challenge database [13] and the Open Access Heart Sound Dataset (OAHS Dataset) [11]. The experiments demonstrate that our proposed method achieves significant improvements over state-of-the-art methods. Under conditions involving Gaussian white noise (GWN) Signal-to-Noise Ratio (SNR) of 0 dB, the proposed method achieves an SNR of 9.01 dB and a Root Mean Square Error (RMSE) of 0.032, outperforming standalone DTCWT and multiple existing models.

The paper is organized as follows. The next section introduces a quick survey of related work of HSS denoising. Section 3 focuses on the systematic processes of a joint HSSs denoising based on DTCWT and ASASS algorithm. The content of Section 4 is the experiment and the results are presented. Finally, a comprehensive summary of the entire text in Section 5.

## 2. Related work

HSS is susceptible to various types of noise during acquisition, such as Electromagnetic Interference (EMI) from the surrounding environment, power frequency interference, bioelectrical interference from the body, breath sounds, and Lung sounds (LS). Denoising methods have been explored across multiple branches of biomedical engineering, including Electrocardiography (ECG), Electroencephalography (EEG), respiratory sounds, and heart sounds, allows to extract from it the maximum amount of efficient and meaningful information.

(i) ECG Denoising. Recent research indicates that the filtering stage should employ an architecture based on Adaptive Filters (AF). Adaptive algorithms enable real-time dynamic optimization by adjusting filter parameters to adapt to input ECG signal characteristics [14]. Additionally, the conventional low-pass and high-pass filters in wavelet transforms have been replaced by fractional-order wavelets. Studies demonstrate through simulations that fractional-order wavelets outperform traditional wavelets in ECG denoising, with their efficacy dependent on wavelet decomposition level selection and coefficient thresholding strategies [15][16]. Deepak H. A. et al. proposes an adaptive thresholding technique combining Empirical Mode Decomposition (EMD) and DTCWT [17]. Traditional Discrete Wavelet Transform (DWT) is prone to induce Gibbs oscillations and frequency aliasing, whereas DTCWT significantly mitigates these issues through enhanced time-frequency decomposition properties [18]. Separate research systematically evaluates the impact of threshold selection, algorithms, and distribution functions on denoising performance within DTCWT frameworks[19]. To address the issue of multi-category noise contamination in real-time acquisition, the Adversarial Denoising Convolutional Network (ADnCNN) was proposed, which enhances model robustness through adversarial training [20]. Enhan Liu et al. introduced the method of Noise Prediction-based ECG Denoising method (NPED) [21]. This method effectively solves noise problems in grayscale images scanned from Paper-based ECGs in hospitals. Furthermore, it overcomes the limitations of traditional waveform redrawing approaches by performing direct denoising to prevent

waveform distortion. For dynamic interferences such as Baseline Wander (BW), Electrode Motion (EM), and Muscular Artifact (MA), the ECG denoising diffusion model (EDDM) was developed to establish a generative diffusion model for signal reconstruction [22]. CNN-SWT integrates the convolutional kernel constraints and architecture of Stationary Wavelet Transform (SWT) into a convolutional neural network (CNN), significantly enhancing the learning efficiency for denoising both linear and nonlinear time-frequency features in ECG signals [23].

(ii) Respiratory sound & EEG Denoising. LS signals are severely contaminated by background noise from multiple sources. Conventional denoising methods may exhibit limited effectiveness due to the non-stationary characteristics of LS and its spectral overlap with various noise sources. A joint denoising approach based on the Butterworth bandpass filter and Sparsity-Assisted Signal Smoothing (SASS) algorithm is proposed, which significantly enhances the signal-to-noise ratio of LS through frequency-domain filtering and sparse-constrained signal reconstruction [24] [25]. An adaptive denoising technique utilizing Discrete Wavelet Transform and Artificial Neural Network (DWT-ANN) has been developed. This method integrates the multi-resolution analysis capability of DWT with the nonlinear adaptive filtering properties of ANN to achieve refined purification of LS signals a noisy environment [26]. In EEG signal processing, Variational Mode Decomposition-based Blind Source Separation (VMD-BSS) and DWT-BSS effectively isolate physiological artifacts while preserving essential neural information [27]. Additionally, researchers have combined EMD with DTCWT to achieve high-fidelity EEG denoising through a two-stage processing framework [28]. Furthermore, an EMD-Hurst analysis combined with spectral subtraction has been implemented to optimize LS denoising through mode selection and energy correction [29]. Besides, a novel deep encoder-decoder-based denoising architecture (LU-Net) has been presented to suppress ambient and internal lung sound noises [11].

(iii) PCG & HSSs Denoising. Digitally acquired heart sound signals via stethoscopes are often distorted by environmental and physiological noise interference, altering their key distinctive characteristics. To address superimposed noise, Xiahou S. et al. employs variational mode decomposition (VMD) for hierarchical filtering [30] [3]. Linear filters have been to utilized to eliminate distortion and interference in fetal phonocardiogram (fPCG) signals [31]. Additionally, integrating Hilbert envelope and homomorphic envelope techniques, combined with reusing pretrained CNN filters for denoising, significantly enhances data utilization efficiency in implantable devices [32]. Incorporating DTCWT with an adaptive neuro-fuzzy inference system (ANFIS) classifier enables simultaneous optimization of signal enhancement and classification [33]. Another study establishes an objective threshold calculation method for DTCWT denoising, optimizing the noise reduction process by quantifying edge preservation and noise elimination as objective functions [34]. In deep learning-based heart sound denoising, Duggan D. et al. proposes a fully convolutional network (FCN) denoising model based on the Spleeter U-Net architecture [10]. For generative adversarial network (GAN)-based models generating normal heart sounds, incorporating EWT denoising effectively reduces GAN training cycles and computational costs [35].

## 3. Methodology

In this paper, a joint HSSs denoising using DTCWT and ASASS algorithm is adopted. Firstly, the original HSS is denoised by Butterworth filter, and then the denoised HSS $X(t)$ is sent to 4-layer DTCWT for decomposition. And then, sparse-assisted signal smoothing (SASS) algorithm, including sparse-assisted signal extraction layer and signal smoothing layer, is used to extract and smooth the decomposed HSS, and adaptive parameter algorithm is set to adjust the SASS parameters. Finally, it is combined with the decomposed signal and the inverse DTCWT transform is used to reconstruct the signal, so as to better realize the denoising of heart sound signal $\hat{X}(t)$. The architecture of our proposed method is shown in Fig. 1.

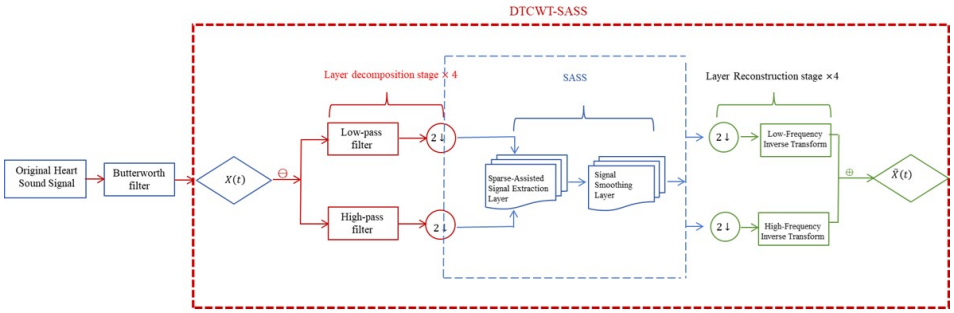


**Fig. 1.** The architecture of the joint HSSs denoising using DTCWT and ASASS

### 3.1. DTCWT

DTCWT is a wavelet transform method possessing time-frequency localization and multiscale characteristics. Compared to traditional single wavelet transforms, it can more effectively capture the time-frequency information of signals. DTCWT decomposes a signal multiscale using a set of orthogonal wavelet basis functions, where the wavelet basis functions at each scale consist of a pair of low-pass and high-pass filters. Unlike traditional wavelet transforms, DTCWT employs two parallel wavelet transform trees, processing the real and imaginary parts of the heart sound signal separately. This structure better preserves the signal phase information and provides superior time-frequency resolution. Fig.2 illustrates the typical structure of DTCWT, encompassing both decomposition and reconstruction stages.

In the decomposition stage, the input signal is fed into two signal processing trees, referred to as the real tree and the imaginary tree. Collectively, these two trees are termed the Complex Wavelet Transform. Each tree can be viewed as a Filter Bank (FB) tree, where the upper tree (Top Tree) contains the real part's low-pass filter ($h_0$) and high-pass filter ($h_1$). The lower tree (Bottom Tree) contains the imaginary part's low-pass filter ($g_0$) and high-pass filter ($g_1$), respectively.

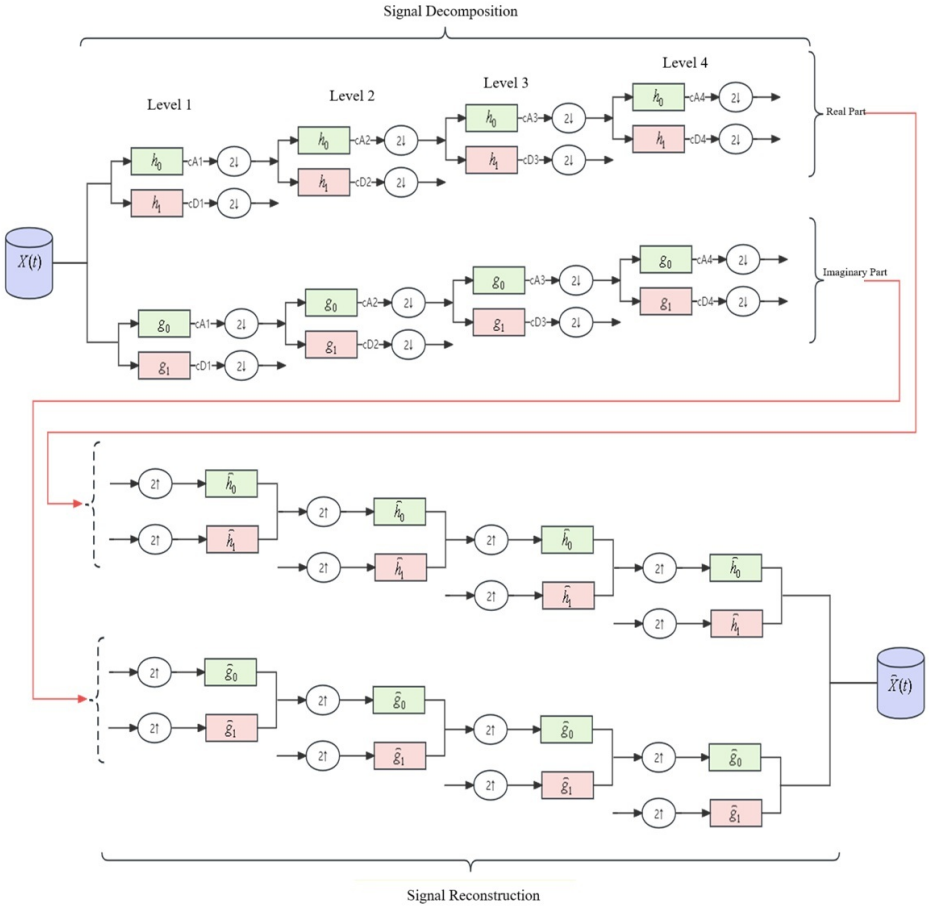


**Fig. 2.** The architecture of DTCWT

At the Level 1 decomposition, the real part is split into low-frequency information ($h_0$) and high-frequency information ($h_1$). The low-frequency coefficients are termed approximation coefficients ($cA_1$), while the high-frequency coefficients are called detail coefficients ($cD_1$). This process is repeated at Level 2, where the previous low-frequency component, $cA_1$, is sub-decomposed into the low-frequency $cA_2$ and the high-frequency $cD_2$.

The decomposition progresses iteratively, concluding at Level 4. Consequently, four approximation coefficients ($cA_1$–$cA_4$) and four detail coefficients ($cD_1$–$cD_4$) are generated for the real part's Discrete Wavelet Transform (DWT). Similarly, during the decomposition stage, this identical process is applied to the imaginary part. This yields a parallel set of four approximation coefficients ($cA_1$–$cA_4$) and four detail coefficients ($cD_1$–$cD_4$) for the imaginary component. This dual-tree structure effectively mitigates limitations inherent in the standard DWT, such as spectral aliasing and lack of shift-invariance. This

process is repeated at Level 2, where the previous low-frequency component, $cA_1$, is sub-decomposed into the low-frequency $cA_2$ and the high-frequency $cD_2$. The decomposition progresses iteratively, concluding at Level 4. Consequently, four approximation coefficients ($cA_1$–$cA_4$) and four detail coefficients ($cD_1$–$cD_4$) are generated for the real part's DWT.

In DTCWT, the given signal is expressed in terms of shifted and dilated forms of the mother wavelet function $\psi_{j,n}(k)$ and a scaling function $\varphi_{j,n}(k)$. $cA_j$ denotes the approximation coefficient at level j and formally expressed as

$$cA_j(n) = \sum_k x(k)\,\psi_{j,n}(k) \tag{1}$$

$$\psi_{j,n}(k) = \frac{1}{\sqrt{2j}}\psi\left(\frac{n - k \bullet 2j}{2j}\right) \tag{2}$$

where $\psi_{j,n}(k)$ represents the mother wavelet function, $k$ is the shift parameter, and $j$ indicates the DTCWT decomposition level.

Similarly, detail coefficients $cD_j(n)$ capturing high-frequency components at decomposition level $j$ and formally expressed as

$$cD_j(n) = \sum_k x(k)\,\varphi(n-k) \tag{3}$$

where $\varphi(n-k)$ represents the mother wavelet function and defined as $\varphi(n) = (-1)^n \psi(N-1-n)$, and $k$ is the shift parameter.

### 3.2. Sparse-assisted Signal Extraction

The fundamental principle of the SASS algorithm leverages the concept of sparse representation, combining linear time-invariant (LTI) filtering with sparse optimization principles to model signals as the sum of a piecewise-smooth component and a low-pass component. By solving the sparse optimization problem, it identifies the sparsest vector that optimally represents the original signal, revealing its primary structure and essential characteristics. The algorithm operates similarly to wavelet denoising but performs denoising via sparse optimization, thereby avoiding the pseudo-Gibbs phenomena commonly encountered in DTCWT.

Let $y(n)$ denote the noisy HSS, which can be represented as

$$y(n) = s_1(n) + s_2(n) + \omega(n), n \in Z \tag{4}$$

where $s_1(n)$ represents a low-frequency signal, $s_2(n)$ has the sparse derivative of M-order, and $\omega(n)$ is additive white Gaussian noise (AWGN). The signal $y$ can be written as

$$y = AQ^{-1}P^TPt + \alpha A^{-1}Q^TQ_1\mu \tag{5}$$

$$C(\mu) = \frac{1}{2}\left|\left|t - A^{-1}P^TPt - \alpha A^{-1}Q^TQ_1\mu\right|\right|_2^2 + \lambda||\mu||_1 \tag{6}$$

where $\mu$ is sparse and is determined by the cost function $C(\mu)$, $P$ and $Q$ are the banded Toeplitz matrices [24].

To achieve enhanced denoising performance, this paper integrates the SASS algorithm with DTCWT. After sub-band decomposition, sparse-assisted signals are extracted from the sub-bands to obtain superior signal characteristics. Through this integration, the sparse-assisted signal is ultimately extracted by solving the following optimization problem

$$s\left(t\right) = arg_{s} \min \left|\left|cD - s\right|\right|_{2}^{2} + \lambda \left|\left|s\right|\right|_{1} \tag{7}$$

Here, $cD$ denotes the detail coefficients obtained via DTCWT decomposition, and $s$ represents the sparse-assisted signal. The final smoothed signal $\hat{X}(t)$ is obtained by weighting and summing the sparse-assisted signal with the decomposed signal as

$$\hat{x}\left(t\right) = x\left(t\right) + \gamma s(t) \tag{8}$$

where $\gamma$ is a smoothing parameter that controls the influence of the sparse-assisted signal during the smoothing process. However, as the four-level DTCWT decomposition through high-pass and low-pass filters yields sub-bands with varying scales, orientations, and threshold requirements, a fixed SASS algorithm cannot uniformly process these heterogeneous sub-bands. Consequently, this paper proposes an adaptive threshold of SASS to further process these sub-band signals. The adaptive strategies dynamically adjust the threshold based on sub-band energy distribution and directional correlations, achieving a balance between noise suppression and feature preservation.

### 3.3. An Adaptive Threshold of SASS Based on Sub-band Energy Distribution

First, the sub-band signals obtained from a four-level DTCWT decomposition are grouped. For each sub-band signal $s_i$, its energy $E_i$ is calculated by summing the squares of its values and can be represented as

$$E_i = \sum_{n=1}^{N} \left(s_i(n)\right)^2 \tag{9}$$

where $N$ is the length of the sub-band signal.

Subsequently, we need to set an initial threshold, adopting the standard deviation threshold method. This is a simple and intuitive approach that uses the standard deviation of the signal as the threshold. It is typically assumed that the noise in the signal follows a Gaussian distribution. Therefore, the standard deviation of the signal can be used to estimate the noise level. The threshold can be chosen as a multiple of the standard deviation with the expression as follows $T = k\sigma$, where $T$ is the threshold, $k$ is the multiplier coefficient, and $\sigma$ is the standard deviation of the signal.

Finally, by comparing the sub-band energy $E_i$ with the threshold $T$, the sub-band signals are divided into a high-energy group and a low-energy group. For each group of sub-band signals, we dynamically adjust the parameters of the SASS method based on their characteristics. As for optimal denoising performance, the SNR of the sub-band signal $s_i$ is used as the metric to design an adaptive algorithm for dynamically tuning the SASS parameters. The SNR is expressed as

$$SNR = 10 \lg \frac{\sum_{n=1}^{N} (x(n))^2}{\sum_{n=1}^{N} (x(n) - \hat{x}(n))^2} \tag{10}$$

Here, $N$ denotes the number of HSS samples, $x(n)$ denotes the original HSS, $\hat{x}(n)$ denotes the denoised HSS.

We continue to set the signal threshold and define the objective function $F$ as the relative change rate of the SNR, specifically the ratio of the post-processing SNR to the pre-processing SNR. Its formula is as

$$F = \frac{SNR_{post} - SNR_{pre}}{SNR_{pre}} \tag{11}$$

where $SNR_{post}$ is the processed SNR, $SNR_{pre}$ is the original SNR.

For each group of sub-band signals, parameters are iteratively updated using the gradient descent method. The parameter update formula is as

$$\theta_{n+1} = \theta_n - \alpha \nabla U(\theta_n) \tag{12}$$

where $\theta_{n+1}$ denotes the updated parameter value, $\theta_n$ represents the current parameter value, and $\alpha$ is the gradient of the objective function and controlling the step size of each iteration. $U(\theta)$ is the objective function, characterizing its dependence on parameters $\theta$. $\nabla U(\theta)$ is the gradient of $U(\theta)$ with respect to $\theta$, indicating the magnitude and direction of change in the objective function at the current parameter value. We set a convergence criterion to stop the iteration when the change in the objective function is less than a predefined threshold, as illustrated by the following

$$|U(\theta_{n+1}) - U(\theta_n)| < \varepsilon \tag{13}$$

where $\varepsilon$ denotes the set convergence decision threshold. The adjusted parameters and the processed sub-band signals are returned. Ultimately, through this adaptive algorithm, we dynamically adjust the parameters of the SASS method based on the characteristics of the sub-band signals, maximizing the relative change rate of the SNR to enhance the denoising effect.

## 4. Experiment

### 4.1. Data Resources

Publicly available heart sound datasets are used to evaluate the performance of our proposed method. (i) PhysioNet Computing in Cardiology Challenge 2016 dataset. All recorded data in this dataset are categorized as normal or pathological (abnormal) samples. We select five imbalanced recording categories (Training-A through E), comprising a total of 3,126 heart sound recordings, with uneven distribution of cardiac condition severity. Additionally, each category contains extraneous emergency noises, such as uncontrolled ambient sounds. We merge Category C and Category D data due to their limited quantity of heart sound signals and excludes significantly distorted signal recordings. 2,735

heart sound recordings are extracted, including 546 pathological recordings and 2,189 normal recordings, representing approximately 87.5% of the entire dataset. (ii) Open Access Heart Sound (OAHS) Dataset. OAHS Dataset is a publicly available cardiac audio database providing researchers with noise-free heart sound data. This dataset contains 1,000 heart sound samples, and the samples are divided into five distinct categories based on cardiac pathological characteristics: Normal (N), Aortic Stenosis (AS), Mitral Regurgitation (MR), Mitral Stenosis (MS), and Mitral Valve Prolapse (MVP). Each category comprises 200 independent representative recordings.

### 4.2. Performance Metrics

In order to evaluate the entire SNR performance, the global error measures include exploited E(SNR), var(SNR) and PE(SNR). E(SNR), var(SNR) and PE(SNR) are the mean value, variance and Percentage error of SNR. Besides, RMSE is another metrics that popular in HSSs analysis. RMSE measures the similarity between the denoised HSS and the original clean HSS by calculating the deviation between them. A smaller RMSE value indicates a smaller difference between the denoised signal and the original clean signal, meaning better denoising performance. RMSE is expressed as

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2} \tag{14}$$

Here, $x_i$ is the $i^{th}$ sample value of the original clean HSS, $\hat{x}_i$ is the $i^{th}$ denoised HSS, and $N$ is the number of signal samples.

### 4.3. Denoising Results

To verify whether our proposed method can perform noise reduction on HSSs in real-world scenarios, we use noisy HSSs from PhysioNet Computing in Cardiology Challenge 2016 dataset, which are polluted by various unavoidable entities. To validate the stability of our proposed method, one being a normal heart sound and the other a pathological heart sound. Fig. 3 and Fig. 4 clearly demonstrates the method's exceptional denoising efficacy on HSSs. For normal heart sounds, it visually preserves the characteristic waveform of S1 and S2 while effectively eliminating interference noise during intervals. Crucially, the model retains pathological features in abnormal heart sounds during denoising, providing essential auxiliary diagnostic information. Notably, post-processing observation reveals significant noise reduction at S1 and S2 locations in original signals. This confirms the model's capability to filter out non-cardiac noise from respiration and muscle movement, resulting in enhanced signal clarity. Importantly, pathological signatures, including murmurs or accentuated heart sounds, remain well-preserved even in cardiac pathology cases. Hence, DTCWT+ASASS can denoise HSSs effectively under real-world noisy environments.

Based on the numerical analysis of E(SNR) and var(SNR) from the Table 1, DTCWT+ASASS exhibits significant superiority in signal processing. DTCWT+ASASS consistently outperforms DTCWT in output E(SNR), var(SNR) and PE(SNR). In the case of Training-A, DTCWT+ASASS provides a E(SNR)(dB) (8.1 vs. 7.9 vs. 7.1), a var(SNR)(dB)

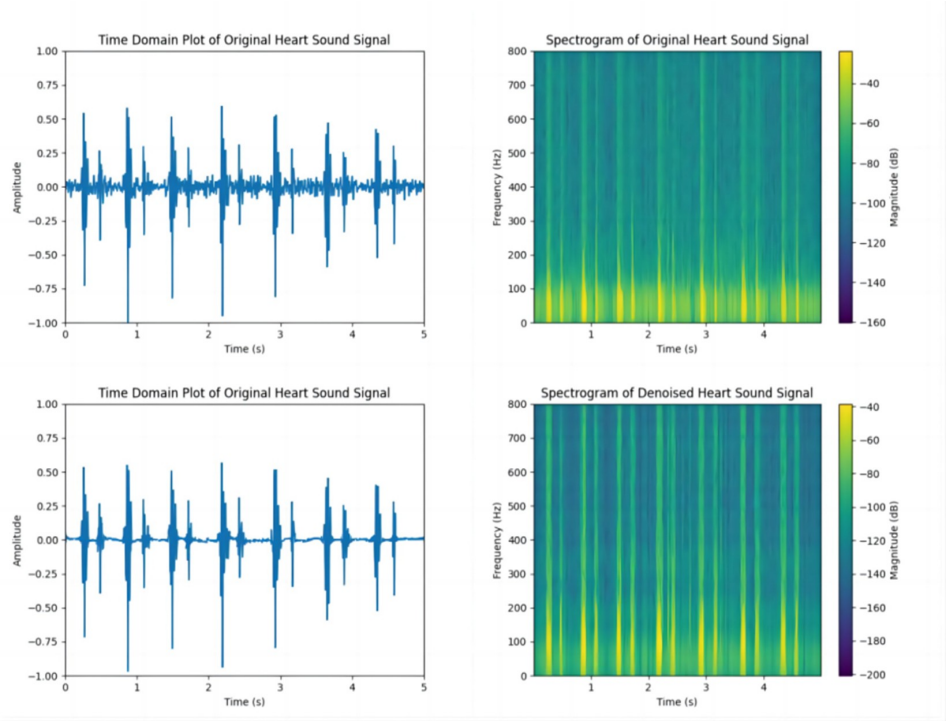


**Fig. 3.** Time-domain and Spectrogram Plots of Normal and Denoised Normal HSSs

(2.2 vs. 2.4 vs. 2.4) and a PE(SNR)(dB) (+0.14 vs. +0.11 vs 0). In particular, a total of 357 samples (255 normal HSSs, 102 abnormal HSSs) from 409 heart sound recordings (292 normal HSSs, 117 abnormal HSSs) were included for analysis in 121 patients with MVP. These samples included HSSs denoised by DTCWT, HSSs denoised by DTCWT+ASASS, and the original HSSs as a control. E(SNR), var(SNR), and PE(SNR) significantly increased with the increase in denoising degree ($p < 0.001$). This indicates that the observed result differences are extremely unlikely to be caused by random variation, further demonstrating the superiority and reliability of DTCWT+ASASS in HSSs denoising.

Fig. 5 and Fig.6 presents (a) Original HSS, (b) Noise-contaminated HSS, (c) HSS denoised by WT, (d) HSS denoised by DTCWT, and (e) HSS denoised by DTCWT+ASASS. WT achieves noise reduction but concurrently introduces significant signal distortion. Although DTCWT mitigates signal distortion, it demonstrates suboptimal denoising performance in overlapping noise-original signal segments and exhibits susceptibility to pseudo-Gibbs phenomena at signal boundaries. DTCWT+ASASS not only delivers superior noise suppression but also preserves clinically relevant acoustic features (e.g., S1/S2 energy retention) while eliminating perceptible distortion. For the denoised normal HSSs, the spectral energy distribution is uniform, with a clear boundary between the S1 and S2 frequency bands, and the energy is highly concentrated. For the denoised pathological HSSs, mitral stenosis leads to the disappearance of the high-frequency components in S1, with

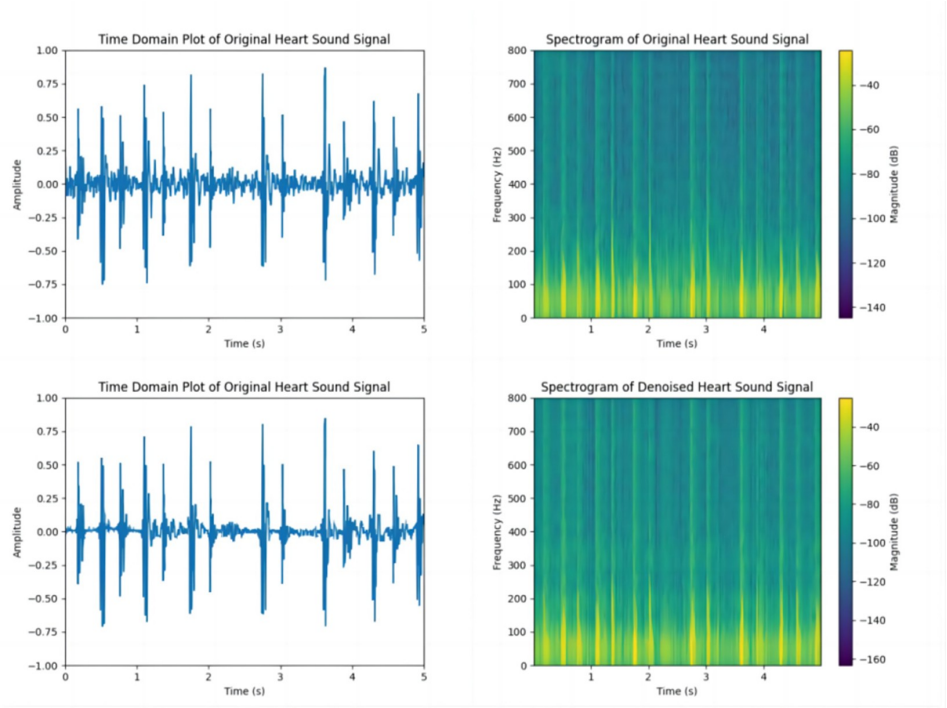


**Fig. 4.** Time-domain and Spectrogram Plots of Pathological and Denoised Pathological HSSs

energy shifting to the low-frequency band, while aortic insufficiency causes an increase in the low-frequency energy in S2. Compared with DTCWT, DTCWT+ASASS can effectively suppresses pseudo-Gibbs artifacts, and outperforms both benchmark methods in denoising efficacy metrics for both normal and pathological heart sounds.

Table 2 compares the denoising performance of multiple algorithms for HSSs contaminated by GWN at different SNRs of -5 dB, 0 dB, 5 dB, 10 dB, 15 dB, and 20 dB. The results show that DTCWT+ASASS consistently outperforms other comparison methods in terms of SNR and RMSE metrics across all noise levels. Similarly, Table 3 compares the denoising effect of HSSs contaminated by pink noise at different noise levels (including -5 dB, 0 dB, and 5 dB). Under conditions involving GWN SNR of 0 dB, DTCWT+ASASS can reduce the RMSE 30.43% and 15.78%, compared to WT and DTCWT. Correspondingly, under conditions involving pink noise SNR of 0 dB, DTCWT+ASASS can reduce the RMSE by 9.89% and 46.42%, compared to WT and DTCWT. Unlike GWN, pink noise has more energy concentrated in the low-frequency range and gradually diminishes in the high-frequency range. In contrast, we select EMD, which is currently commonly used, as the comparison method for HSSs denoising. This table also shows that DTCWT+ASASS has greater advantages over EMD and DTCWT in terms of SNR and RMSE metrics. Overall, compared to WT, EMD, and DTCWT, DTCWT+ASASS stronger anti-noise ability and greater stability.

**Table 1.** Denoising performance comparison among DTCWT, DTCWT+ASASS

| Recording Categories | Methods | E(SNR) | var(SNR) | PE(SNR) |
|---|---|---|---|---|
| Training-A | Original HSS | 7.1 | 2.4 | 0 |
| | DTCWT | 7.9 | 2.4 | +0.11 |
| | DTCWT+ASASS | 8.1 | 2.2 | +0.14 |
| Training-B | Original HSS | 0.7 | 0.6 | 0 |
| | DTCWT | 10.8 | 3.5 | +13.96 |
| | DTCWT+ASASS | 11.1 | 3.5 | +14.86 |
| Training-C+D | Original HSS | 5.6 | 3.7 | 0 |
| | DTCWT | 8.9 | 4.0 | +0.58 |
| | DTCWT+ASASS | 9.5 | 3.8 | +0.69 |
| Training-E | Original HSS | 5.6 | 3.0 | 0 |
| | DTCWT | 9.3 | 4.8 | +0.65 |
| | DTCWT+ASASS | 10.5 | 3.9 | +0.88 |

In this way, DTCWT+ASASS has the following characteristics:

(i) The DTCWT achieves time-shift invariance through its unique dual-tree structure, enabling more effective handling of transient features in non-stationary signals while preserving time-frequency localization properties. Compared to conventional WT, the DTCWT demonstrates superior noise separation capability during HSS denoising, thereby enhancing reconstruction accuracy.

(ii) The ASASS algorithm is an advanced HSS processing technique primarily used to smooth high-frequency signals and suppress overlapping noise frequencies in selected bands, thereby recovering noise-free signals. By effectively capturing fine features within HSS fluctuations, the DTCWT+ASASS method achieves superior SNR compared to the DTCWT.

**Table 2.** Comparison of Denoising Performance at Gaussian White Noise Levels

| Noise Level | Performance metrics | WT | DTCWT | DTCWT+ASASS |
|---|---|---|---|---|
| -5 dB | SNR | 4.78 | 4.66 | 5.65 |
| | RMSE | 0.064 | 0.052 | 0.050 |
| 0 dB | SNR | 7.53 | 8.62 | 9.01 |
| | RMSE | 0.046 | 0.038 | 0.032 |
| 5 dB | SNR | 11.76 | 11.48 | 13.23 |
| | RMSE | 0.028 | 0.026 | 0.024 |
| 10 dB | SNR | 16.12 | 18.01 | 18.18 |
| | RMSE | 0.017 | 0.016 | 0.013 |
| 15 dB | SNR | 19.01 | 20.96 | 21.52 |
| | RMSE | 0.013 | 0.010 | 0.008 |
| 20 dB | SNR | 23.88 | 24.37 | 25.68 |
| | RMSE | 0.007 | 0.006 | 0.005 |

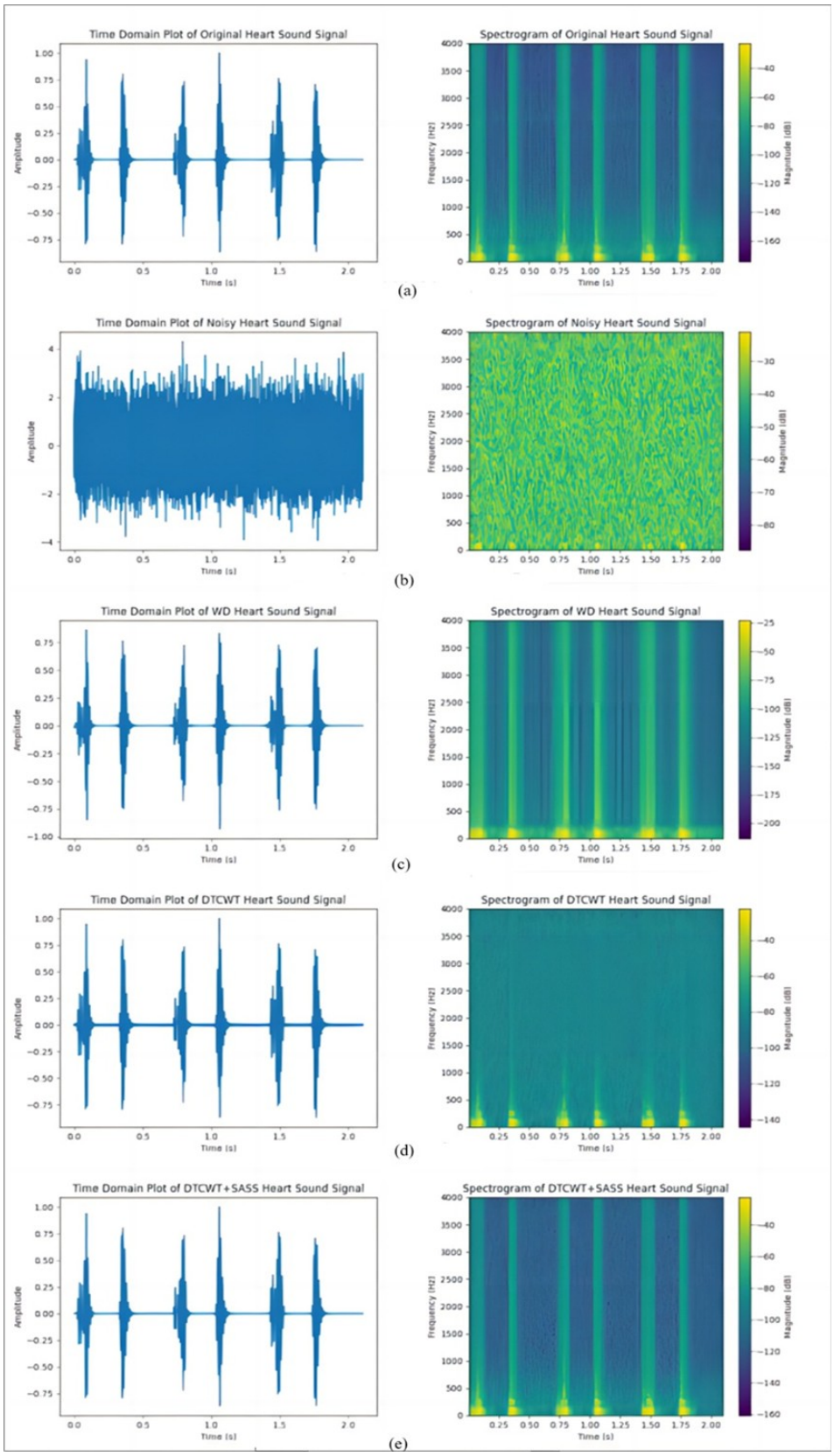


**Fig. 5.** Denoising Effects of Three Methods on Normal HSSs under Gaussian White Noise with SNR=0 dB: (a) Original HSS, (b) Noise-contaminated HSS, (c) HSS denoised by WT, (d) HSS denoised by DTCWT, and (e) HSS denoised by DTCWT+ASASS

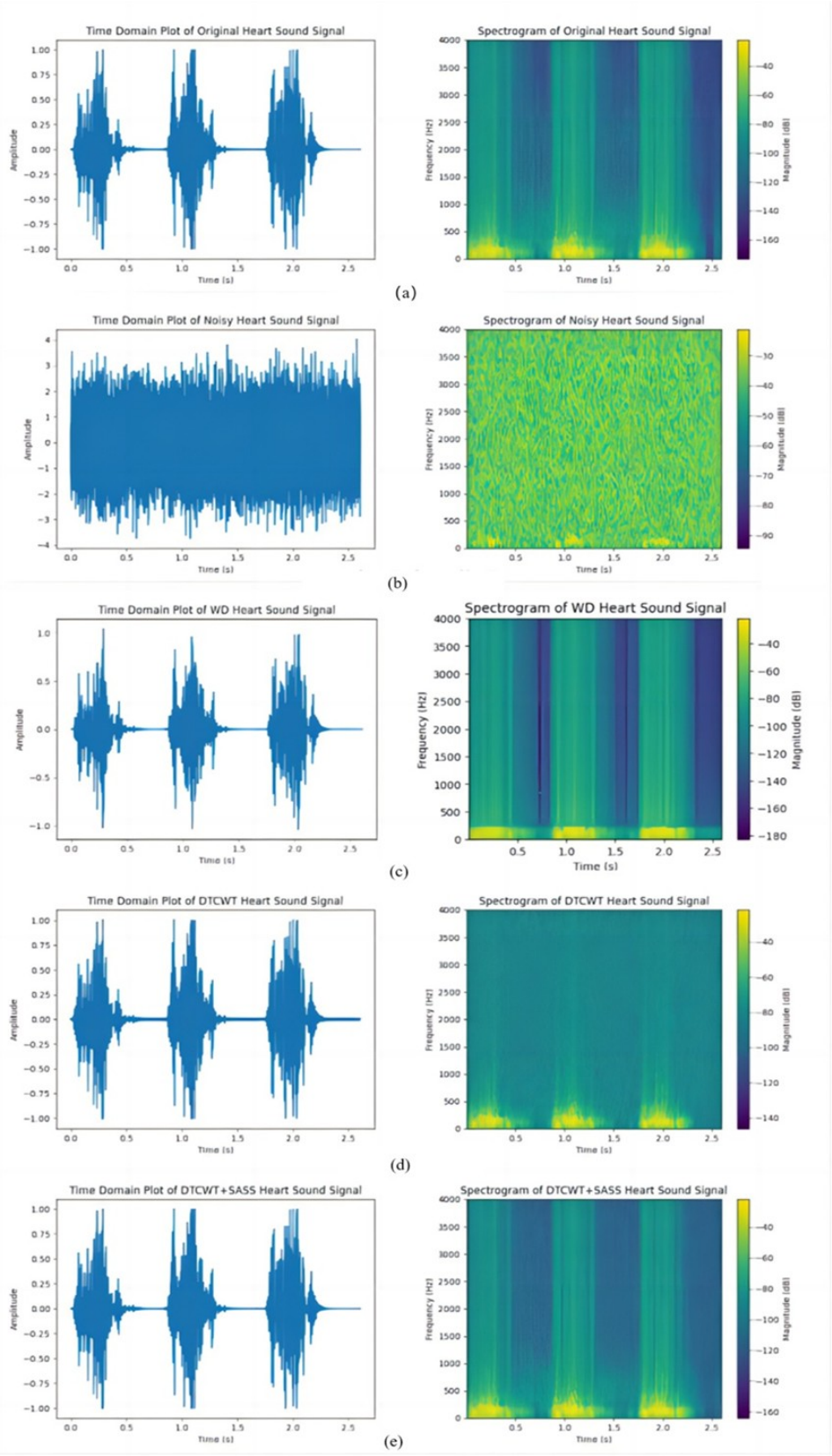


**Fig. 6.** Denoising Effects of Three Methods on Aortic Stenosis Murmur under Gaussian White Noise with SNR=0 dB: (a) Original HSS, (b) Noise-contaminated HSS, (c) HSS denoised by WT, (d) HSS denoised by DTCWT, and (e) HSS denoised by DTCWT+ASASS

**Table 3.** Comparison of Denoising Performance at Pink Noise Levels

| Noise Level | Performance metrics | EMD | DTCWT | DTCWT+ASASS |
|---|---|---|---|---|
| -5 dB | SNR | 3.25 | 2.82 | 4.43 |
| | RMSE | 0.162 | 0.188 | 0.101 |
| 0 dB | SNR | 1.62 | 2.63 | 5.62 |
| | RMSE | 0.093 | 0.084 | 0.045 |
| 5 dB | SNR | 6.51 | 7.62 | 11.33 |
| | RMSE | 0.053 | 0.043 | 0.023 |

## 5. Conclusion

In this work, we propose a joint HSSs denoising method using DTCWT and ASASS. First, HSS undergoes preprocessing via Butterworth filtering to reduce high-frequency noise. Subsequently, the signal is decomposed using DTCWT to obtain its multi-scale feature representation. The decomposed signal is then fed into the ASASS algorithm, which automatically adjusts its parameters through a designed adaptive mechanism. Leveraging its sparse adaptive characteristics, the DTCWT+ASASS model extracts critical feature information from the signal and reconstructs it. This approach addresses the limitations of conventional denoising methods, such as signal distortion and low SNRs, while simultaneously overcoming the inherent pseudo-Gibbs phenomena of DTCWT at signal boundaries. Besides, we have validated it on PhysioNet/CinC 2016 database and OAHS Dataset, and both healthy and pathological recordings. Both normal and aortic stenosis HSSs contaminated with AWGN or pink noise at different levels of SNR, our approach achieves significant improvements over state-of-the-art methods. Under conditions involving GWN SNR of 0 dB, the proposed method achieves an SNR of 9.01 dB and a RMSE of 0.032, outperforming standalone DTCWT and multiple existing models.

**Acknowledgment.** This work was supported by Natural Science Foundation of Fujian Province under Grant No. 2023J011426; Engineering Research Center for Big Data Application in Private Health Medicine of Fujian Universities, Putian University, Putian, Fujian 351100, China (MKF202408).

## References

1. Hadiyoso, S., Mardiyah, D., Ramadan, D., Ibrahim, A.: Implementation of electronic stethoscope for online remote monitoring with mobile application. Bulletin of Electrical Engineering and Informatics 9(4), 1595–1603 (2020)
2. Centracchio, J., Parlato, S., Esposito, D., Andreozzi, E.: Accurate localization of first and second heart sounds via template matching in force cardiography signals. Sensors 24(5), 1525 (2024)
3. Hu, J., Hu, Q., Liang, M.: Heart sounds classification using adaptive wavelet threshold and 1d ldcnn. Computer Science and Information Systems 20(4), 1483–1501 (2023)
4. Xiao, F., Liu, H., Lu, J.: A new approach based on a 1d+2d convolutional neural network and evolving fuzzy system for diagnosis of cardiovascular disease from heart sound signals. Applied Acoustics 216, 109723 (2024)

5. Gilles, J.: Empirical wavelet transform. IEEE Transactions on Signal Processing 61(16), 3999–4010 (2013)
6. Al-Shannaq, M., Nasrawi, A., Bsoul, A., Saifan, A.: Abnormal heart sound recognition using svm and lstm models in real-time mode. Scientific Reports 15(1), 9129 (2025)
7. Nawaz, S., Li, J., Li, D., Shoukat, M., Bhatti, U., Raza, M.: Medical image zero watermarking algorithm based on dual-tree complex wavelet transform, alexnet and discrete cosine transform. Applied Soft Computing 169, 112556 (2025)
8. Nia, P., Hesar, H.: Abnormal heart sound detection using time-frequency analysis and machine learning techniques. Biomedical Signal Processing and Control 90, 105899 (2024)
9. Li, J., Ke, L., Du, Q.: Classification of heart sounds based on the wavelet fractal and twin support vector machine. Entropy 21(5), 472 (2019)
10. Duggan, D., Temko, A., Sarana, V., Factor, A., Popovici, E.: Denoising of heart sounds using lightweight fcns and spectrograms with and without context. IEEE Access (2025)
11. Ali, S., Shuvo, S., Al-Manzo, M., Hasan, A., Hasan, T.: An end-to-end deep learning framework for real-time denoising of heart sounds for cardiac disease detection in unseen noise. IEEE Access 11, 87887–87901 (2023)
12. Hu, J., Chen, L., Yang, M., Shen, S., Gao, X.: Psbd-ewt-egan: Heart sound denoising using psbd-ewt and enhancement generative adversarial network. Computer Science and Information Systems 00, 5–5 (2025)
13. Liu, C., Springer, D., Li, Q., Moody, B., Juan, R., Chorro, F., Castells, F., Roig, J., Silva, I., Johnson, A.: An open access database for the evaluation of heart sound algorithms. Physiological Measurement 37(12), 2181 (2016)
14. Mat Rozi, N., Hashim, F., Shaharuddin, S., Miskan, M., Ahmad, K., Salleh, M.: Comparison on wavelet adaptive filter performance in denoising ecg signal (2023)
15. Houamed, I., Saidi, L., Srairi, F.: Ecg signal denoising by fractional wavelet transform thresholding. Research on Biomedical Engineering 36(3), 349–360 (2020)
16. Hu, J., Chen, L., Shen, S., Wang, T.: Explainable multi-agent deep reinforcement learning for joint task offloading and resource allocation in distance- and channel-aware noma vehicular edge networks. IEEE Internet of Things Journal (2025)
17. Deepak, H., Vijayakumar, T.: Optimal threshold estimation using grey wolf optimization for emd-dtcwt based ecg denoising. International Journal of Recent Technology and Engineering (2020)
18. Wang, F., Ji, Z.: Application of the dual-tree complex wavelet transform in biomedical signal denoising. Bio-Medical Materials and Engineering 24(1), 109–115 (2014)
19. Prashar, N., Sood, M., Jain, S.: Design and implementation of a robust noise removal system in ecg signals using dual-tree complex wavelet transform. Biomedical Signal Processing and Control 63, 102212 (2021)
20. Hou, Y., Liu, R., Shu, M., Chen, C.: An ecg denoising method based on adversarial denoising convolutional neural network. Biomedical Signal Processing and Control 84, 104964 (2023)
21. Liu, E., Liu, Y., Wang, Z., Zhou, H., Wang, X.: A noise prediction method for paper-based grayscale ecg denoising. Biomedical Signal Processing and Control 111, 108408 (2026)
22. Li, Z., Tian, Y., Jin, Y., Wei, X., Wang, M., Liu, J., Liu, C.: Eddm: A novel ecg denoising method using dual-path diffusion model. IEEE Transactions on Instrumentation and Measurement (2025)
23. Peng, H., Chang, X., Yao, Z., Shi, D., Chen, Y.: A deep learning framework for ecg denoising and classification. Biomedical Signal Processing and Control 94, 106441 (2024)
24. Haider, N., Behera, A.: Respiratory sound denoising using sparsity-assisted signal smoothing algorithm. Biocybernetics and Biomedical Engineering 42(2), 481–493 (2022)
25. Haider, N., Behera, A.: Computerized respiratory sound-based diagnosis of pneumonia. Medical & Biological Engineering & Computing 62(1), 95–106 (2024)

26. Pouyani, M., Vali, M., Ghasemi, M.: Lung sound signal denoising using discrete wavelet transform and artificial neural network. Biomedical Signal Processing and Control 72, 103329 (2022)
27. Massar, H., Drissi, T., Nsiri, B., Miyara, M.: Advancements in blind source separation for eeg artifact removal. Applied Acoustics 228, 110300 (2025)
28. Anupallavi, S., Ashokkumar, S., Premkumar, M., Sangeetha, R.: Enhanced eeg signal classification through integrated dtcwt, human learning optimization, and optimized dbn-hmm models for improved performance. Biomedical Signal Processing and Control 110, 108229 (2025)
29. Haider, N.: Respiratory sound denoising using empirical mode decomposition, hurst analysis and spectral subtraction. Biomedical Signal Processing and Control 64, 102313 (2021)
30. Xiahou, S., Liang, Y., Ma, M., Du, M.: A strong anti-noise segmentation algorithm based on variational mode decomposition and multi-wavelet for wearable heart sound acquisition system. Review of Scientific Instruments 93(5) (2022)
31. Faradisa, I., Ananda, A., Sardjono, T., Purnomo, M.: Denoising of fetal phonocardiogram signal by wavelet transformation. In: E3S Web of Conferences. vol. 188, p. 00013. EDP Sciences (2020)
32. Alali, S., Kachenoura, A., Albera, L., Hernandez, A., Michel, C., Senhadji, L., Karfoul, A.: Optimized cnn-based denoising strategy for enhancing longitudinal monitoring of heart failure. Computers in Biology and Medicine 184, 109430 (2025)
33. Al-Naami, B., Fraihat, H., Al-Nabulsi, J., Gharaibeh, N., Visconti, P., Al-Hinnawi, A.: Assessment of dual-tree complex wavelet transform to improve snr in collaboration with neuro-fuzzy system for heart-sound identification. Electronics 11(6), 938 (2022)
34. Renuka, S., Edla, D.: Adaptive shrinkage on dual-tree complex wavelet transform for denoising real-time mr images. Biocybernetics and Biomedical Engineering 39(1), 133–147 (2019)
35. Narváez, P., Percybrooks, W.: Synthesis of normal heart sounds using generative adversarial networks and empirical wavelet transform. Applied Sciences 10(19), 7003 (2020)

**Jianqiang Hu** is an associate professor in school of computer and information engineering, Xiamen University of Technology, China. He once worked as a postdoctoral researcher at Tsinghua University. He received his Ph.D. degree in computer science and engineering from National University of Defense Technology, China, in 2005. He is the author of more than 70 articles, including respected journals such as CHINESE JOURNAL OF COMPUTERS, IEEE INTERNET OF THINGS JOURNAL and IEEE TRANSACTIONS ON BIG DATA. His current research interests include Edge Computing, Biomedical Signal Processing, and Big Data Analytics.

**Dafeng Shen** is a master student at school of computer and information engineering, Xiamen University of Technology, China. Her research interests include Edge Computing, Biomedical Signal Processing and Deep Reinforcement Learning.

**Lin Chen** is a master student at school of computer and information engineering, Xiamen University of Technology, China. His research interest include Edge Computing and Biomedical Signal Processing.

**Yu Chen** is a master student at school of computer and information engineering, Xiamen University of Technology, China. His research interests include Biomedical Signal Processing and Deep Reinforcement Learning.

**Shigen Shen** received the B.S. degree in fundamental mathematics from Zhejiang Normal University, Jinhua, China, in 1995, the M.S. degree in computer science and technology from Zhejiang University, Hangzhou, China, in 2005, and the Ph.D. degree in pattern recognition and intelligent systems from Donghua University, Shanghai, China, in 2013. He is a Professor with the School of Information Engineering, Huzhou University, Huzhou, China. He has published more than 100 technical papers, including respected journals such as IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and IEEE INTERNET OF THINGS JOURNAL. His current research interests include Internet of Things, cyber security, edge computing, and game theory.

**Yan Che** is a professor at Putian University. She received her ME degree in Computer Application from Xiamen University, Xiamen, China, in 2005, and Ph.D. degree in Control Theory and Engineering from Donghua University, Shanghai, China, in 2016. She is currently working in Engineering Research Center for Big Data Application in Private Health Medicine of Fujian Universities, Putian University, Putian, Fujian, China. Her research interests include Data Processing, Data Analysis and Data Security.

*Received: August 28, 2025; Accepted: January 5, 2026.*