Sparse ECG Denoising with Generalized Minimax Concave Penalty

Zhongyi Jin [1], Anming Dong [2,*] , Minglei Shu [1,*] and Yinglong Wang [1]

[1] Shandong Provincial Key Laboratory of Computer Networks, Shandong Computer Science Center (National Supercomputer Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Jinan 250014, China; zyjin.qlut@gmail.com (Z.J.); wangylscsc@126.com (Y.W.)

[2] School of Computer Science and Technology, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250353, China

* Correspondence: anmingdong@qlu.edu.cn (A.D.); shuml@sdas.org (M.S.)

check for
updates

Abstract: The electrocardiogram (ECG) is an important diagnostic tool for cardiovascular diseases. However, ECG signals are susceptible to noise, which may degenerate waveform and cause misdiagnosis. In this paper, the ECG noise reduction techniques based on sparse recovery are investigated. A novel sparse ECG denoising framework combining low-pass filtering and sparsity recovery is proposed. Two sparsity recovery algorithms are developed based on the traditional ℓ_1-norm penalty and the novel generalized minimax concave (GMC) penalty, respectively. Compared with the ℓ_1-norm penalty, the non-differentiable non-convex GMC penalty has the potential to strongly promote sparsity while maintaining the convexity of the cost function. Moreover, the GMC punishes large values less severely than ℓ_1-norm, which is utilized to overcome the drawback of underestimating the high-amplitude components for the ℓ_1-norm penalty. The proposed methods are evaluated on ECG signals from the MIT-BIH Arrhythmia database. The results show that underestimating problem is overcome by the proposed GMC-based method. The GMC-based method shows significant improvement with respect to the average of output signal-to-noise ratio improvement (SNR_{imp}), the average of root mean square error (RMSE) and the percent root mean square difference (PRD) over almost any given SNR compared with the classical methods, thus providing promising approaches for ECG denoising.

Keywords: ECG denoising; sparse recovery; Generalized Minimax Concave Penalty (GMC); ℓ_1-norm

1. Introduction

As the recording of the bioelectric potentials produced by rhythmical cardiac activities, the Electrocardiogram (ECG) signals provides plentiful and valuable information on the heart functional conditions, and it is thus has been widely used not only for the clinical diagnosis of cardiovascular diseases in hospital, but also for patient monitoring at home. However, the ECG signals are essentially weak and non-stationary, and so susceptible to various noises. The noise may blur or even mask the key characteristics in the ECG waveform, which will further degenerate the reliability and efficiency of the clinical diagnosis. Therefore, it is important and valuable to denoise the ECG, i.e., extract high resolution cardiac signals from noisy ECG observations.

ECG noise elimination is a very challenging task due to the spectral overlap of signals and noise, and a wide range of ECG denoising techniques have been continuously and extensively studied in the last few decades. These methods can be typically classified as the linear filtering, empirical mode decomposition (EMD) and wavelet transform, etc. They are the most well-developed approaches and have their own advantages and drawbacks.

The linear filtering methods [1–4] have low computational complexity and fast calculation speed and thus have been widely used in embedded or portable monitoring instruments. However, the classical filters failed to reduce the noise components that overlap with signal spectrum. Moreover, they will cause significant damage to the details of ECG signals which contain sharp edges and impulses of short duration. Therefore, they are often used in combination with other noise reduction algorithms.

To overcome these difficulties, nonlinear methods have been investigated and Wavelet-based methods are the most famous and effective [5–14]. These approaches rely on the idea of thresholding under the assumption that signal magnitudes will concentrate in the Wavelet representation while the noise will spread, such that the noise can be illuminated by cutting the Wavelets coefficients less than a pre-determined threshold. The performance of the Wavelet denoising methods depends on the matching degree of the selected Wavelet basis functions and the ECG signals. The basis functions will be fixed once the algorithms have been developed and may not necessarily match the practical ECG signals.

Due to the ability of dealing with non-stationary and nonlinear signals, the EMD-based ECG denoising methods gained extensive attention recently [15–21]. The EMD method was first introduced in [22] for analyzing data from nonstationary and nonlinear processes. The EMD decomposes a given signal into a series of intrinsic mode functions (IMFs) through an iterative process called sifting. The IMFs play the roles of basis functions, and the EMD method can be seen as a type of Wavelet decomposition with adaptive basis functions. Since the basis functions of the EMD are derived from the signal itself, it has the potential to outperform the traditional Wavelet methods. However, the EMD has a limitation of mode mixing, which causes the mixing of higher order components with the lower order components. Because of the mode mixing problem, it is difficult to separate the signal and noise components clearly even in higher order IMFs. To overcome this problem, paper [21] proposed a grey spectral noise cancellation (GSNC) scheme for ECG signals where two-stage discrimination is employed with the empirical mode decomposition (EMD), the ensemble empirical mode decomposition (EEMD) and the grey spectral noise estimation (GSNE). This method effectively alleviates the mode mixing problem. In the case of denoising, removing some of the IMFs will inevitably corrupt the useful information of the signal [23]. For example, in ECG denoising, the EMD may cause damage to P-waves and T-waves, which may result in misdiagnosis [24].

Either the Wavelet or the EMD methods can be classified as the transform-domain processing approaches under the assumption that the true signal can be well approximated by a linear combination of some basis functions [25,26]. The essence underlying these methods is that the signal can be sparsely represented in the transform domain. Since the ECG signals consist of periodical pulse components, they can also be well represented by a combination of low pass components and sparse components in time domain. Consequently, ECG denoising methods based on the time-domain sparse representations have been an active research area in recent years. Paper [27] modeled the ECG signals as the sum of two second and third order sparse derivative signals to reduce noise. In [28], an sparse ECG signals enhancement and QRS waves detection algorithm was proposed. In [29,30], the sparse optimizationis combined with linear time invariant (LTI) filtering to conduct ECG signals denoising. In these sparse representation methods, ℓ_1-norm is used as a penalty term to construct sparse optimization models. This penalty term has the ability to control the sparsity of the optimization results. However, the ℓ_1-norm penalty has a drawback that it always underestimates the interested signal [31], since larger coefficients are penalized more heavily in the ℓ_1 norm than smaller coefficients [32]. The underestimating causes the amplitude of estimated signals to be lower than the true value, which may lead to misdiagnoses.

In order to improve the denoising performance, a variety of non-convex penalty terms have been developed to replace ℓ_1-norm [32–34]. Paper [32] proposes a weighted formulation of ℓ_1-norm minimization designed to more closer to the solution of ℓ_0-norm. In [33], the author uses $\ell_2 - \ell_0$ penalty instead of ℓ_1-norm penalty. Paper [34] addresses the minimization problem of ℓ_0-regularized least squares for a continuum and propose two heuristic search algorithms. However, these methods

adopt non-convex penalty to replace ℓ_1-norm, which leads to nonconvex objective functions. As a result, the algorithms become complicated and the solution is often stuck in the local optimal solutions. To overcome these drawbacks, a novel multivariate non-separable non-convex penalty, namely generalized minimax-concave (GMC) penalty, has been recently presented in [35,36]. The GMC penalty is a non-differentiable non-convex penalty obtained by subtracting a differentiable generalized Huber function from ℓ_1-norm. It strongly promotes sparsity while maintaining the convexity of the cost function.

Currently, the GMC penalty has been successfully applied to the fault detection of gearbox and achieved good results [31,37]. However, the application of GMC penalty to the ECG denoising has not been well studied yet. It is not known whether the GMC penalty can be applied to the ECG signals processing area to improve the denoising performance. This motivates the work in this paper to explore the GMC-based sparse denoising method.

In this paper, we study the sparse denoising techniques for the ECG signals. Through modeling the observed ECG signals as a sum of low-pass component, sparse component and additive noise component, we propose an ECG denoising framework by combining the LTI low-pass filtering and the sparsity recovering. As the key ingredients of the framework, the sparsity recovery algorithms are investigated and we proposed two sparsity recovery algorithms based on ℓ_1-norm penalty and GMC penalty, respectively. We compare the denoising performance of the proposed two schemes and demonstrate that the GMC-based scheme preservers more detail information with respect to QRS waves of ECG signals. The contributions of this work are summarized as follows.

1. A novel ECG denoising framework based on sparsity recovery is proposed. Low-pass filtering and sparse denoising idea was first proposed in [38], but it is not known if it can be applied to the ECG denoising. Moreover, the sparsity in [38] deals with the derivatives of the signal, while we directly deal with the sparsity components of the signal. Therefore, we propose a different sparse framework for ECG denoising.
2. We extend the application of GMC penalty to the area of ECG denoising. Although the GMC penalty has been proposed in [35] and several application examples were given there, it is not known how to apply it in the ECG area and whether the performance of ECG denoising can be improved.
3. The performance of the GMC-based sparse denoising algorithm is comprehensively tested. We carry out a series of experiments to explore the performance the GMC-based denoising algorithm. The relationship between the parameter setting and the performance are demonstrated through experiments. It is also shown that the GMC-based algorithm helps to overcome the underestimation problem and outperforms the classical methods known in the literature.

The paper is organized as follows. We introduce the sparsity recovery denoising framework in Section 2. The ℓ_1-norm based algorithm is described in Section 3. The GMC-based sparse denoising algorithm is proposed in Section 4. Experimental results are presented in Section 5. Section 6 concludes the work.

2. System Model

In this section, the sparse signal model in time domain is presented. A sparse denoising framework based on this signal model is established with the help of mathematical derivations.

2.1. Signal Model

The noisy ECG signals can be modeled as the superposition of a low-pass signal, a sparse signal and the additive white Gaussian noise (AWGN), which can be written as

$$y(n) = l(n) + d(n) + w(n), \quad n \in \mathbb{Z} \tag{1}$$

where n is the index to the discrete sampling sequences, $l(n)$, $d(n)$ and $w(n)$ represent the low-pass signal, the sparse signal and the additive white Gaussian noise (AWGN), respectively. The 'sparse' property means that there are many zeros in a signal frame with N data points. To easily understand, we can regard $d(n), 0, 1, 2, \ldots, N-1$, as the periodic components in the ECG signal frame. In the following, to facilitate our derivations, we omit the time index (n) in the case of not being confused.

2.2. Denoising Framework

We aim to recover the pure ECG signals $l + d$ from the observed noisy signals s. To achieve this, a straightforward method is the traditional linear filtering procedure. Specifically, let $s \triangleq l + d$ represent the pure ECG signals, then it can be approximately estimated through traditional linear filtering

$$LPF(y) = LPF(l + d + w) , \tag{2a}$$

$$= LPF(l + w) + LPF(d) , \tag{2b}$$

$$= \acute{l} + LPF(d) , \tag{2c}$$

where LPF represents a low-pass filtering (LPF) process and $\acute{l} \triangleq LPF(l + w)$. Here, we use the zero-phase filter since it can reduce noise in the signals and preserves the QRS complex at the same time, it occurs in the original [38,39]. To facilitate readers, the design method of the LPF is sketched in Appendix A.

From Equation (2c), it is observed that the traditional LPF will smooth not only the low-pass component but also the sparse component in the original signals. Since the sparse component d consists of high frequency ingredients, it is prone to be damaged by the LPF, which will be embodied by the injury of the details of the ECG signals. For example, the amplitude of the peaks, the shape of the Q and R waves may be damaged by the LPF.

As mentioned above, the traditional LPF methods may cause damage to the details of the ECG signals due to the high-frequency ingredients. In order to enhance the details in the recovered signals, we propose a new ECG denoising framework as shown by Figure 1. The main idea is to recover these details, which is sparse, from the residual of the low-pass filter. Specifically, the residual of the LPF is written as

$$y - LPF(y) = HPF(y) , \tag{3a}$$

$$= HPF(l + d + w) , \tag{3b}$$

$$= HPF(d) + HPF(l + w) , \tag{3c}$$

$$= HPF(d) + \tilde{w} , \tag{3d}$$

where HPF represents a high-pass filter (HPF). It is assumed that the LPF and HPF are compensative, i.e., $HPF = I - LPF$. Here, I represents the identity operator. From Equation (3d), we know that the residual of the LPF consists of the high-passed ingredient of the sparse signals plus high-frequency noise. If $\tilde{d} \triangleq HPF(d)$ is recovered, then it can be combined with the output of the LPF ($\tilde{l} = \acute{l} + LPF(d)$) as shown in Figure 1. The combined signal can be written as

$$\tilde{s} = \tilde{l} + \tilde{d} , \tag{4a}$$

$$\approx \acute{l} + LPF(d) + HPF(d) . \tag{4b}$$

Recall that $HPF(d)$ contains the details of the sparse high-frequency components, the details of the recovered ECG signals can be thus improved. Then the problem boils down to how to recover this high-frequency sparse signals, i.e., $HPF(d)$, from noise.

In the following, we propose two methods to recover the sparse components in the residual of the LPF. The first one is based on the ℓ_1-norm penalty, which improves the details in the detains of

the output of LPF but still has obvious underestimating phenomenon. The second one is based on the GMC penalty, which achieves better performance than the ℓ_1-norm method and overcomes the underestimating problem.
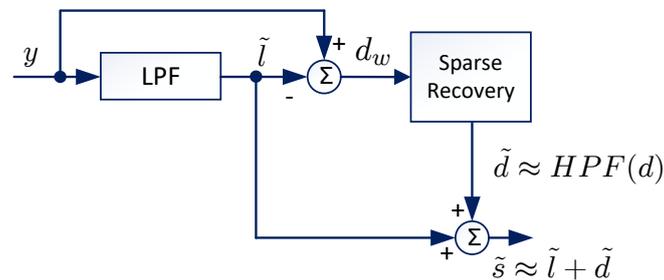

Figure 1. Block diagram of the sparsity electrocardiogram (ECG) denoising process.

3. Sparsity Recovery with ℓ_1-Norm

In this section, we propose the ℓ_1-norm based sparsity recovery algorithm for the sparsity recovery block in Figure 1.

From Figure 1, we know that the performance of the sparsity recovery plays the vital role in the whole denoising procedure. The classical sparsity recovery methods can be applied here straightforwardly, e.g., the ℓ_1-norm penalized least squares method [40]. Specifically, the sparse signal $\tilde{d}$ can be decomposed as

$$\tilde{d} = Ax \,, \tag{5}$$

where $A \in \mathbb{R}^{N \times M}$ is the basis vectors of the signal space and $x \in \mathbb{R}^M$ is the representation coefficients. Given A, we expect to find a sparse solution x to satisfy the equality. The sparse solving process can be mathematically reformulated as the following ℓ_1-norm optimization problem

$$\min_x \quad \|x\|_1 \tag{6}$$
$$s.t.: \quad \tilde{d} = Ax \,.$$

Note that A can be constructed using different methods in principle [35]. In this work, we construct A based on the Short Time Fourier Transform (STFT) matrix as shown in Appendix B, considering that the ECG signals are periodic pulses in time series.

When $\tilde{d}$ is corrupted by noise, i.e., $d_w \triangleq \tilde{d} + \tilde{w}$ is observed, an approximation solution to Equation (6) can be found by solving the following basis pursuit denoising (BPD) problem [41,42]

$$\arg\min_x \frac{1}{2} \|d_w - Ax\|_2^2 + \lambda \|x\|_1 \,. \tag{7}$$

Specifically, the forward-backward splitting (FBS) method can be applied to find the solution of Equation (7) [43], as listed in Algorithm 1. It is noted that the $soft(\cdot)$ in the algorithm is the soft thresholding operator defined by

$$soft(t, T) = t * max\{0, 1 - T/|t|\} \,, \tag{8}$$

where $T \in R_+$ is a given threshold. The $maxeig(\cdot)$ denotes the maximum eigenvalue of a matrix.

Our simulation results show that the details in the recovered ECG signals of the proposed ℓ_1-norm method are enhanced significantly compared with the output of LPF. However, there still exists an obvious gap between the recovered signal and the noiseless signal with respect to the peaks. That is to say the proposed ℓ_1-norm method underestimates the peaks values in the signals. This drawback of the ℓ_1-norm methods has also been revealed in literature [31,35]. To further improve the performance

of the denoised signals with respect to the details, we propose a non-convex sparse regularization method based on the generalized minimax concave (GMC) penalty in the following.

Algorithm 1 FBS algorithm for ℓ_1-norm penalized least squares problem

---

Input: $x^{(0)}, A, A^T, \lambda$
Output: $x^{(i+1)}$
　1: **Set** $\rho = maxeig(\|A^T A\|)$
　2: **Set** $\mu : 0 < \mu < 1/\rho$
　3: Number of iteration: N_{iter}
　4: **for** $i = 0$ *to* N_{iter} **do**
　5:　　$w^{(i)} = x^{(i)} - \mu A^T (Ax^{(i)} - d_w)$
　6:　　$x^{(i+1)} = soft(w^{(i)}, \mu\lambda)$
　7: **end for**
　8: **return** $x^{(i+1)}$

---

4. Sparsity Recovery with GMC Penaly

As mentioned above, the classical sparsity recovery method based on ℓ_1-norm penalty will lead to underestimation to the peaks values. In this section, instead of the ℓ_1-norm penalty, we adopt a novel penalty for the sparsity recovery problem, which can keep the convexity of the least squares cost function minimized and avoids underestimation the values of waveform.

4.1. GMC Penalty

The GMC penalty function $\psi_B : \mathbb{R}^N \to \mathbb{R}$ is defined as

$$\psi_B(x) = \|x\|_1 - S_B(x) \tag{9}$$
$$= \|x\|_1 - \min_{v \in \mathbb{R}^N} \left\{ \|v\|_1 + \frac{1}{2} \|B(x - v)\|_2^2 \right\},$$

where, $B \in \mathbb{R}^{M \times N}$ and $S_B \triangleq \inf_{v \in \mathbb{R}^N} \{\|v\|_1 + \frac{1}{2}\|B(x - v)\|_2^2\}$ is the generalized Huber function [35], $\inf\{\cdot\}$ denotes the infimum of a function.

Figure 2 illustrates the scalar version of GMC penalty functions. It shows that the GMC penalty has a basic property that large values are penalized no less than small values, but the degree of punishment is not as severe as ℓ_1-norm [35]. Because of this property, it becomes possible to overcome the drawback of underestimating the high-amplitude components for the ℓ_1-norm penalty. Moreover, the GMC penalty function is non-convex itself, but it can be carefully customized to maintain the convexity of the cost function to be minimized. This means that the global optimization of the cost function can be achieved with GMC penalty.
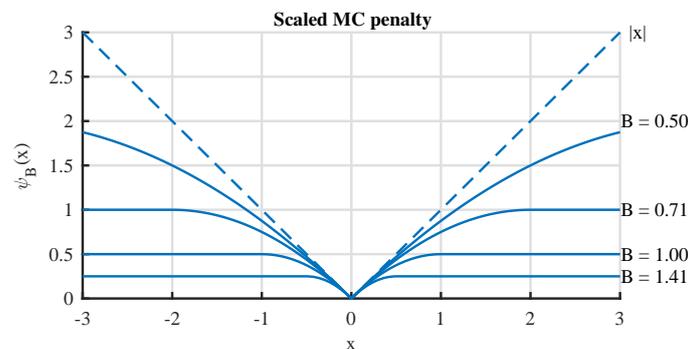

Figure 2. Illustration of the generalized minimax concave (GMC) penalty function.

4.2. Problem Formulation

By replacing the ℓ_1-norm penalty in Equation (1) with the GMC penalty Equation (9), we present the following sparsity recovery problem

$$\arg \min_x \{\frac{1}{2}\|d_w - Ax\|_2^2 + \lambda \psi_B(x)\} . \tag{10}$$

In order to maintain the convexity of the objective function, the parameters in B should be set carefully based on the following lemma.

Lemma 1. *In [35], Let $y \in \mathbb{R}^M$, $A \in \mathbb{R}^{M \times N}$ and $\lambda > 0$, the function $F(x) = \frac{1}{2}\|y - Ax\|_2^2 + \lambda \psi_B(x)$ is a convex function if $B^T B \preceq \frac{1}{\lambda} A^T A$. ($B^T B \preceq \frac{1}{\lambda} A^T A$ implies that $A^T A - \lambda B^T B$ is positive semidefinite.)*

Based on Lemma 1 and given A, the convexity of the objective in Equation (10) can be satisfied by simply setting

$$B = \sqrt{\gamma/\lambda} A, 0 \leq \gamma \leq 1 . \tag{11}$$

- The non-convexity of the penalty ψ_B is controlled by the parameter γ. When $\gamma = 0$, the elements of B are zeros, and the penalty degenerates to the ℓ_1-norm. When $\gamma > 0$, a higher γ will lead to higher non-convexity.

4.3. Proposed GMC-Based Sparsity Recovery Algorithm

Similar with the classical ℓ_1-norm penalty problem, there exists no closed-form solution for the GMC penalty problem Equation (10). In fact, the gradient of the GMC penalty is also inexplicit. In order to solve it, we substitute Equation (9) into Equation (10) and obtain a reformulated optimization problem

$$(x^{opt}, v^{opt}) = \arg \min_x \max_v F(x, v) \tag{12}$$

where $F(x, v) = \frac{1}{2}\|d_w - Ax\|_2^2 + \lambda\|x\|_1 - \lambda\|v\|_1 - \frac{\gamma}{2}\|A(x - v)\|_2^2$ with $0 \leq \gamma \leq 1$, and $d_w = \tilde{d} + \tilde{w}$.

Problem Equation (12) is a saddle-point optimization problem [44], which can also be solved by the FBS method [35,45]. We list details of the solving process in Algorithm 2.

Based on the sparse denoising framework in Figure 1 and the sparsity recovery algorithm in Algorithm 2, we propose the GMC-penalty ECG denoising method. The detailed processing flow of the method is shown in Figure 3 with some key parameter settings. The whole process consists of four steps. At the first step, the noisy ECG signal is sent to a low-pass filter. At the second step, the output of the LPF $\tilde{l}$ and the residual signal $y - \tilde{l}$ are obtained. Two parameters of the LPF, i.e., the order k and the cutoff frequency is w_c should be configured beforehand. The next step is to perform sparse recovery following Algorithm 2. The linear transform matrix A, the data length R, the data overlapping ration and the window function should be set before applying the algorithm. Moreover, the regularization parameter λ and the convexity-guaranteeing factor γ are also set at the start of the algorithm. After processing d_w by Algorithm 2, the optimal solution x^{opt} is obtained for each data frame. Since x^{opt} is the sparse coefficients in frequency domain, we should further recover the denoised sparse signal through the calculation $\tilde{d} = A^H x^{opt}$. Finally, we obtain the denoised ECG signal by synthesizing the low-pass component $\tilde{l}$ and the high-frequency sparse component $\tilde{d}$ in the last step.
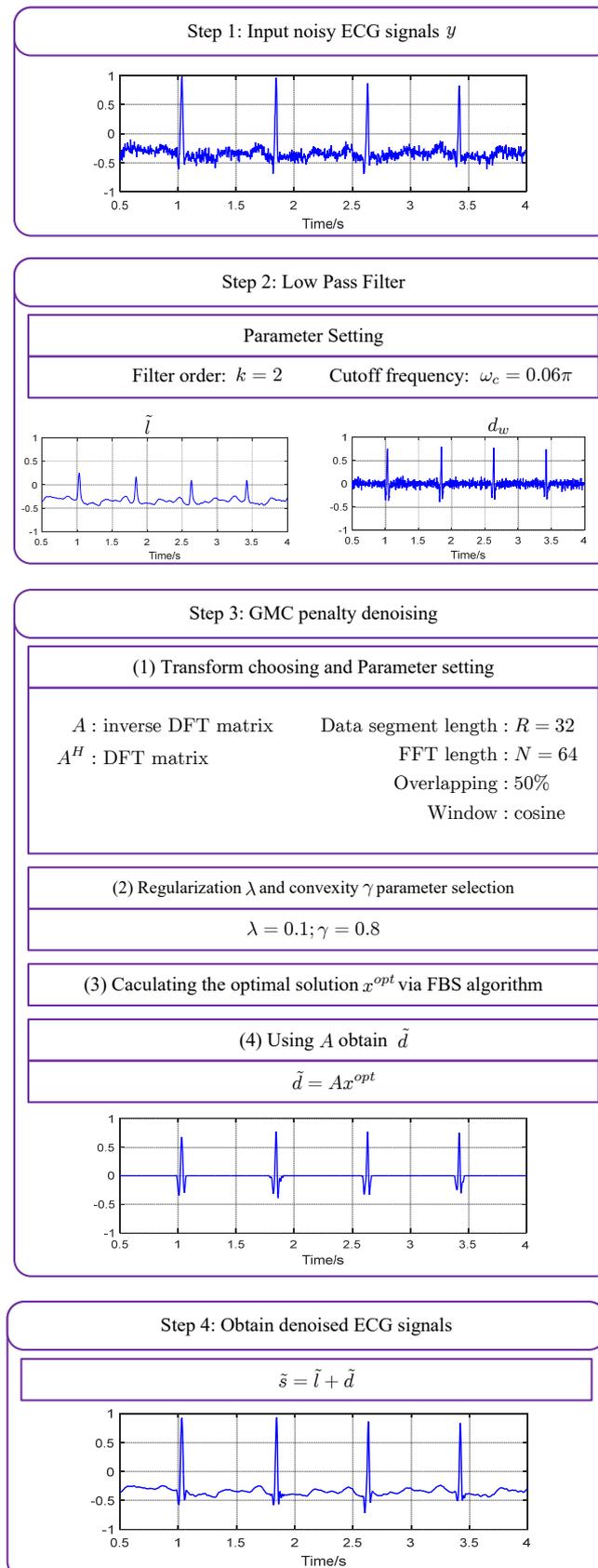
Step 1: Input noisy ECG signals y

Step 2: Low Pass Filter

Parameter Setting

Filter order: $k = 2$ Cutoff frequency: $\omega_c = 0.06\pi$

$\tilde{l}$ d_w

Step 3: GMC penalty denoising

(1) Transform choosing and Parameter setting

A : inverse DFT matrix   Data segment length : $R = 32$
A^H : DFT matrix   FFT length : $N = 64$
                         Overlapping : 50%
                         Window : cosine

(2) Regularization λ and convexity γ parameter selection

$\lambda = 0.1; \gamma = 0.8$

(3) Caculating the optimal solution x^{opt} via FBS algorithm

(4) Using A obtain $\tilde{d}$

$\tilde{d} = Ax^{opt}$

Step 4: Obtain denoised ECG signals

$\tilde{s} = \tilde{l} + \tilde{d}$

Figure 3. The detail flowchart of sparse ECG denoising with GMC-penalty.

Algorithm 2 FBS algorithm for GMC penalized least squares problem

---

Input: $x^{(0)}, v^{(0)}, A, A^T, \lambda$
Output: $x^{(i+1)}$

  1: **Set** $\rho = max\{1, \gamma/(1-\gamma)\} \cdot maxeig(A^T A)$
  2: **Set** $\mu : 0 < \mu < 2/\rho$
  3: Number of iteration: N_{iter}
  4: **for** $i = 0$ *to* N_{iter} **do**
  5:      $w^{(i)} = x^{(i)} - \mu A^T(Ax^{(i)} - d_w) + \mu\gamma A^T A(v^{(i)} - x^{(i)})$
  6:      $u^{(i)} = v(i) - \mu\gamma A^T A(v^{(i)} - x^{(i)})$
  7:      $x^{(i+1)} = soft(w^{(i)}, \mu\lambda)$
  8:      $v^{(i+1)} = soft(u^{(i)}, \mu\lambda)$
  9: **end for**
10: **return** $x^{(i+1)}$

---

5. Simulation Results

5.1. Configurations and Preprocessing

The performance of the proposed algorithms are evaluated based on the ECG records of the MIT-BIH arrhythmia database [46]. The database contains a total of 48 records, each record length is 30 min, the sampling rate is 360 Hz. We note that the original signals in the MIT-BIH database have been filtered by a band-pass filter with a passband from 0.1 to 100 Hz. However, it is obvious that the noise components falling in the passband will not be eliminated. To understand its influence on the clean signal, we make an observation of the original signals in the database, as shown in Figure 4a. It can be observed that there exist noise components in the signals, but the amplitudes of the noise components are relatively weak compared with the useful signal. In order to gain deeper insight into the influence of the noise on our algorithms, we process the original signal directly using our algorithms, and the results are shown in Figure 4b. It can be perceived that the weak noise in the original signal is 'smoothed' out by our algorithms. That is to say our algorithms have the ability to suppress the noise in the acquired signals. However, considering that the original noise in the original signals of the database is relatively weak, we omit its impact on the 'true' signal. In this case, we call the original signals in the MIT-BIH database as 'pure signal', despite there existing very weak noise in practice.
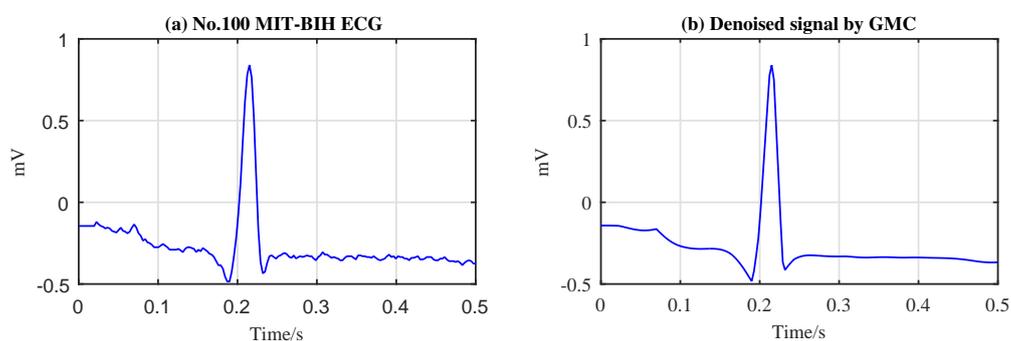

Figure 4. Denoising effect of original signal of MIT-BIH database.

The performance of the proposed algorithm is compared with the works in [30,47,48]. Three different performance criteria are chosen for the performance measurement, i.e., the rooted mean square error ($RMSE$), the percent root mean square difference (PRD) an the signal-to-noise ratio improvement (SNR_{imp}).

For the LPF, we set the parameter of filter order as $k = 2$. The the digital cutoff frequency is set as $w_c = 0.06\pi$, i.e., $f_c = 0.03 * F_s = 10.8$ Hz, where $F_s = 360$ Hz is the sampling frequency of the ECG signals. It is noted that the cutoff frequency should be customized for different ECG signals, since the periodic characteristics of the ECG signals are not uniform in practice. We took data from the first 1 min, a total of 21,600 points. The data segment length R is set as 32, and the frame length of the STFT is 64.

To illustrate the proposed denoising framework, as shown by Figure 1, we plot the outputs of different preprocessing nodes in Figure 5. Specifically, Figure 5a is the original ECG signal read from the No.100 recode in the MIT-BIH database. This pure signal is then corrupted by AWGN noise with 10dB SNR and the result waveform is shown in Figure 5b. The noisy ECG signal, i.e., s, is the input of our denoising framework. It is sent to the LPF, and the output of which is shown in Figure 5c. By substituting this LPF signal from the raw noisy signal, a residual signal is obtained which can be treated approximately as a residual high-frequency sparse component of the ECG signal contaminated with noise, as shown in Figure 5d.
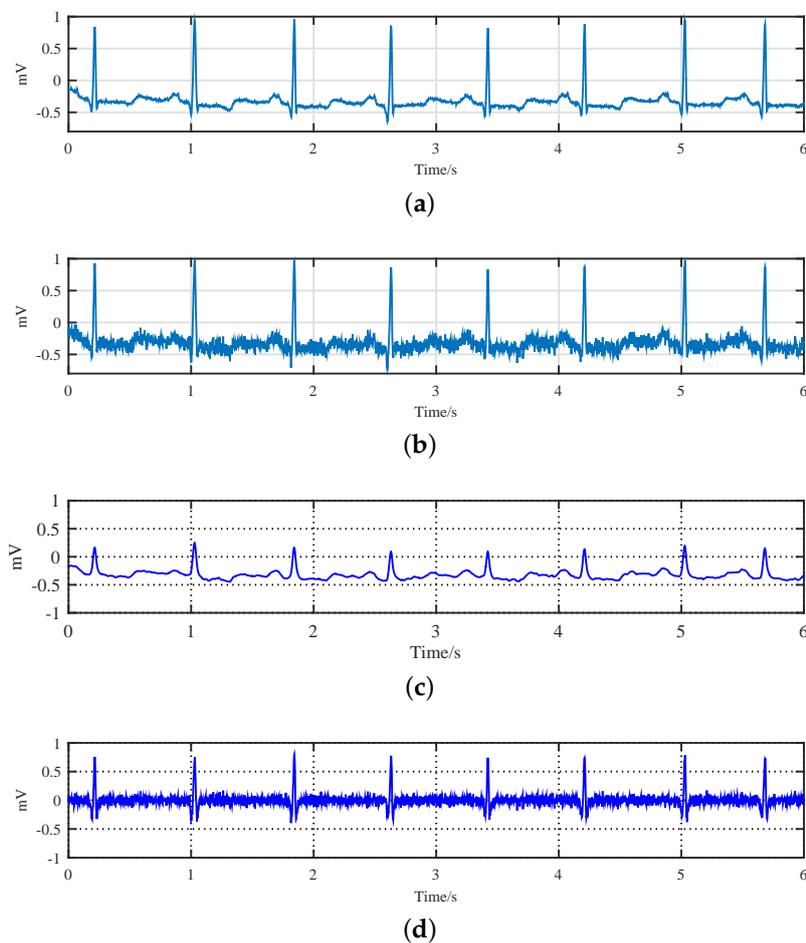

Figure 5. Comparison of the original ECG signal y, the noisy ECG signal s, the low-pass filtered signal $\tilde{l}$ and the residual sparse component d_w in the proposed sparse denoising framework. (**a**) No.100 MIT-BIH ECG; (**b**) Noisy ECG with 10 dB SNR; (**c**) Output of the LPF; (**d**) Residual noisy sparse component.

5.2. Convergence of the Proposed Sparsity Recovery Algorithms

The noisy sparse signal $d_w = \tilde{d} + \tilde{w}$ is further sent to the sparsity recovery block in Figure 1, where the proposed ℓ_1-norm method or the GMC-penalty method is applied to extract the sparse signal. The convergence curves of the algorithms are plotted in Figure 6. In the simulation, we set $\gamma = 0.8$

and $\lambda = 0.1$, and we define the update rate as $\frac{\tilde{d}(k)-\tilde{d}(k-1)}{\tilde{d}(k)}$, where k denotes the kth iteration. It can be observed that the updating rate of the algorithm will be less than 0.1% after about 250 iterations. This means the proposed algorithm will converge after hundreds of iterations. In Figure 6, we also plot the converge performance of the SASS method which is proposed in [30]. It can be observed that the update rate of the SASS algorithm will be less than 10^{-4} after 250 steps, which reveals that the proposed algorithms in this work converge slower than the SASS algorithm.
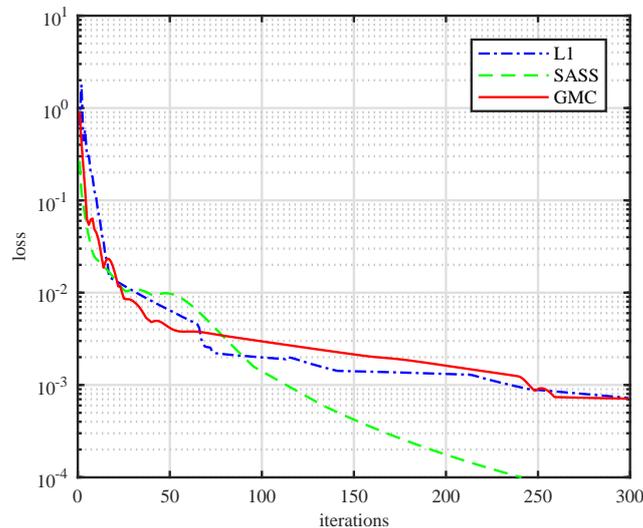
Figure 6. The updating rate of the recovered signal versus iterating numbers.

5.3. Performance Comparison of the Proposed Algorithms

The outputs of the sparsity recovery block are plotted in Figure 7. It can be observed that the noise in the d_w (i.e., Figure 5d is removed by both algorithms).

As the final step of the proposed denoising scheme, we obtain the final denoised ECG signal by adding the outputs of the LPF and the sparsity recovery block according to Figure 1. The resulting waveforms are shown in Figure 8. It can be observed that all the three methods, i.e., the proposed ℓ_1-norm method, the proposed GMC method and the SASS method, eliminate the noise effectively. However, the ℓ_1-norm method underestimates the R-peaks visibly as shown in the down-right corner of Figure 8. The SASS method performs better than the ℓ_1-norm method with respect to the R-peaks, but it underestimates the Q-peaks as shown in the down-left corner of Figure 8. As a comparison, the GMC method outperforms the other two methods with respect to both the R-peaks and Q-peaks.

To gain more insights into the proposed algorithms, we plot the time-frequency relationships of the recovered ECG signals in Figure 9. It is observed that the GMC-based method preserves more high-frequency components than the ℓ_1-norm method, which can be perceived in the spectrogram.
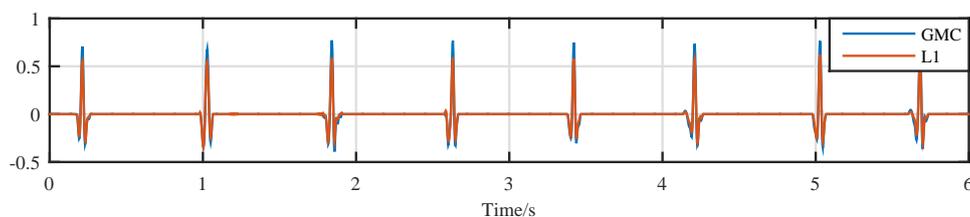

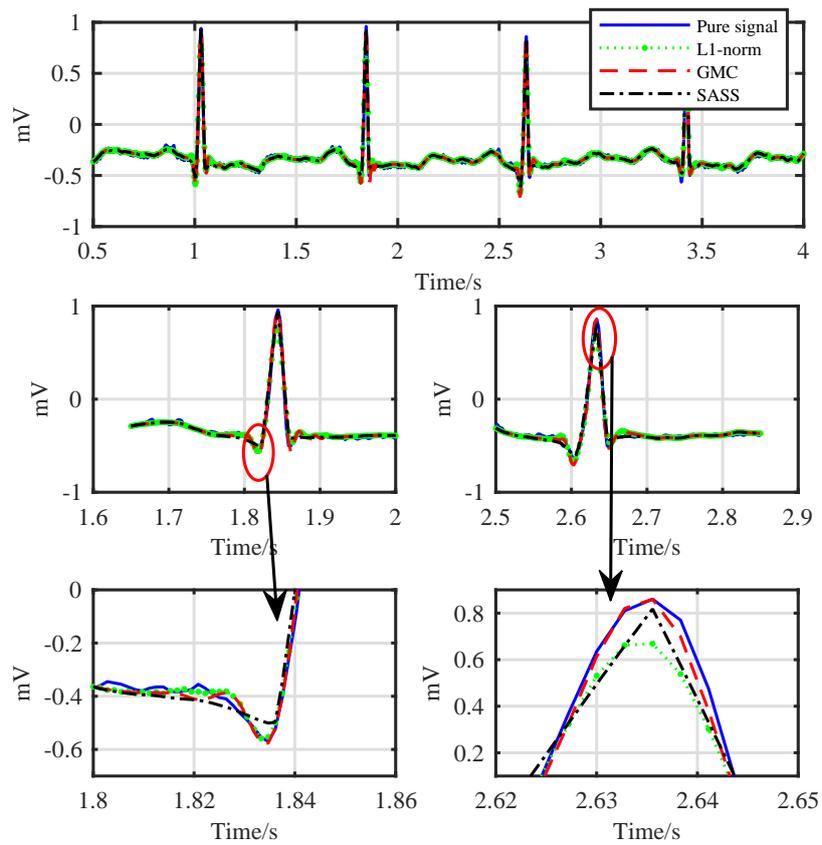**Figure 7.** Output of the sparsity recovery block.

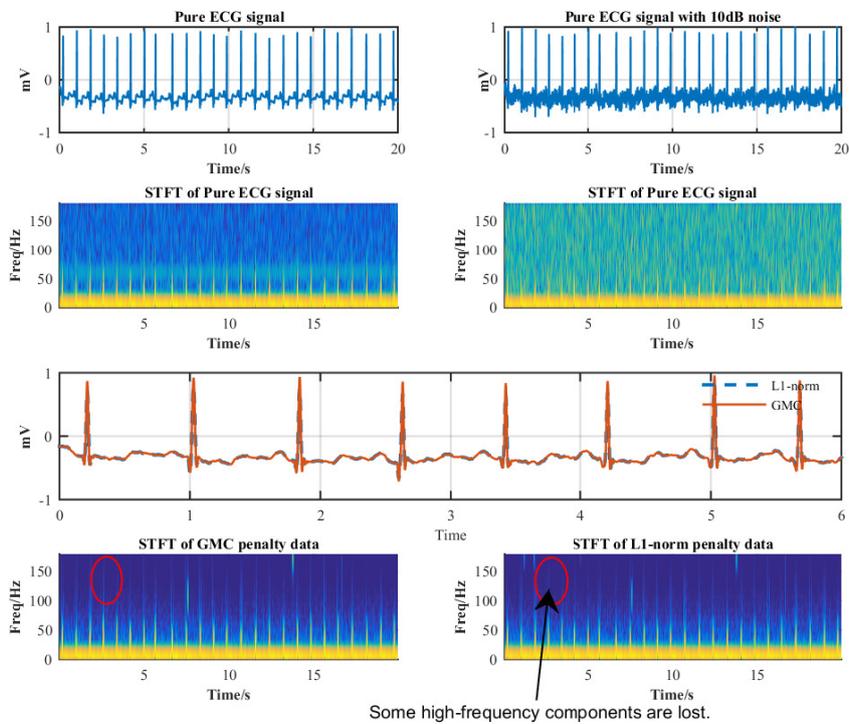**Figure 8.** Comparison of the recovered signals in time domain.

Figure 9. Comparison time-frequency properties of the GMC and ℓ_1-norm methods.

5.4. Impacting of λ

For the GMC-based algorithm, both λ and γ should be specified. Since the two parameters are coupled, we can fix one of them and vary the other to study their impact on the performance of the method. Specifically, we set $\gamma = 0.8$, and varying λ from 0 to 0.6. The resulted RMSE performance of the algorithm is shown in Figure 10. It can be observed that a better RMSE can be achieved when λ is chosen in the vicinity of 0.09.
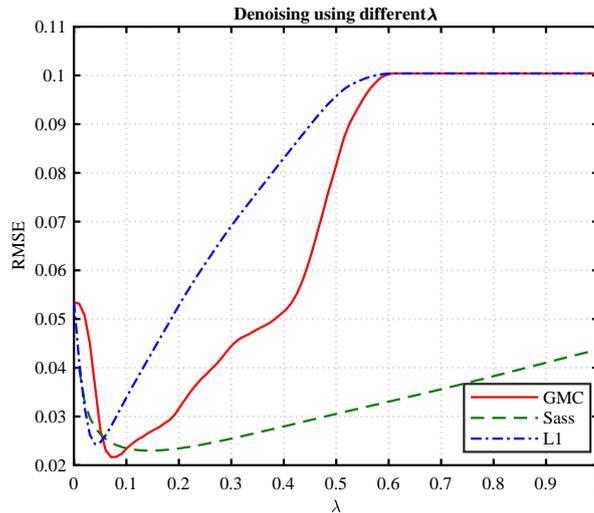

Figure 10. RMSE versus λ with $\gamma = 0.8$.

In Figure 11, we compare the rooted mean square error ($RMSE$), the percent root mean square difference (PRD) and the signal-to-noise ratio improvement (SNR_{imp}) achieved by different methods over different SNRs (20 dB 15 dB 10 dB 5 dB -5 dB -10 dB). The SASS [30], NLM [47] and CE-EMD-AN [48] algorithms are adopted as references. At each SNR, we average the results of all 48 data in the MIT-BIH database. It can be observed that the proposed GMC-based algorithm outperforms the referenced algorithms for all SNRs.
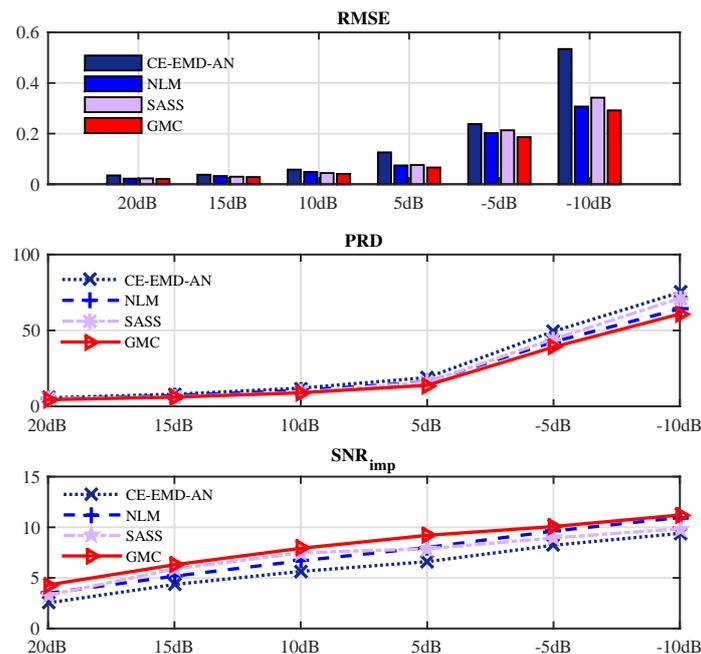

Figure 11. Performance evaluation over RMSE, PRD and SNR_{imp} criteria.

To fully demonstrate the performance of the proposed algorithm, we list the SNR_{imp} values achieved by different algorithms over SNR = 10 dB in Table 1. The results show that the GMC-based algorithm performs the best of all of the experimented SNRs.

Table 1. Comparison of the SNR_{imp} achieved by different methods with input SNR = 10 dB.

Dateset	GMC	SASS	NLM	CEEMDAN	Dateset	GMC	SASS	NLM	CEEMDAN
mitdb/100	7.656	7.015	6.843	6.046	mitdb/201	8.212	7.505	6.619	6.361
mitdb/101	7.616	7.458	6.567	6.029	mitdb/202	8.530	8.121	7.704	7.596
mitdb/102	6.745	4.713	5.949	5.341	mitdb/203	5.805	5.336	4.537	3.448
mitdb/103	9.083	7.800	7.597	5.675	mitdb/205	8.105	7.196	6.236	5.088
mitdb/104	7.079	6.375	6.559	5.776	mitdb/207	9.252	8.541	8.224	6.721
mitdb/105	8.576	8.267	7.073	3.828	mitdb/208	7.838	8.382	6.324	5.050
mitdb/106	7.912	7.382	5.867	4.992	mitdb/209	7.632	7.276	6.085	4.457
mitdb/107	7.925	8.188	7.450	6.923	mitdb/210	8.183	8.083	7.170	4.965
mitdb/108	6.962	6.921	6.297	4.352	mitdb/212	7.166	7.042	5.938	4.612
mitdb/109	8.326	8.394	7.565	7.559	mitdb/213	8.217	8.158	7.134	4.872
mitdb/111	7.162	6.761	6.200	5.476	mitdb/214	8.410	8.292	7.462	4.706
mitdb/112	8.318	7.826	6.682	5.680	mitdb/215	7.168	7.181	6.116	5.443
mitdb/113	8.668	8.098	7.353	5.224	mitbd/217	7.652	7.688	6.116	5.645
mitdb/114	6.102	6.032	5.371	4.885	mitdb/219	8.667	7.963	7.723	4.377
mitdb/115	9.417	7.948	7.503	6.675	mitdb/220	8.767	7.586	6.653	6.977
mitdb/116	7.825	7.434	6.823	3.651	mitdb/221	8.916	7.632	6.670	5.163
mitdb/117	8.122	8.136	6.506	7.973	mitdb/222	6.596	6.566	5.909	5.200
mitdb/118	7.366	5.043	6.078	5.774	mitdb/223	8.580	7.920	7.540	7.163
mitdb/119	8.575	7.842	7.504	6.274	mitdb/228	6.777	7.549	6.752	6.039
mitdb/121	9.702	8.525	7.598	4.631	mitdb/230	8.209	7.860	6.376	5.844
mitdb/122	7.856	7.401	7.048	7.105	mitdb/231	8.076	7.859	7.225	6.152
mitdb/123	9.007	7.692	7.037	6.093	mitdb/232	2.206	4.824	3.741	3.666
mitdb/124	9.252	8.164	7.374	6.858	mitdb/233	8.973	8.894	6.500	6.305
mitdb/200	8.190	8.134	6.638	5.842	mitdb/234	8.265	8.193	7.014	6.284

In order to more intuitively show the denoising performance of each algorithm, we illustrate in Figure 12 the values of RMSE, PRD and SNR_{imp} achieved by the algorithms for the data number as 100, 101, 103, 105, 106, 115 and 220 in the database. It is noted that these data are selected because they were used in literature [47]. In fact, our algorithm can be applied to any ECG data in the database.
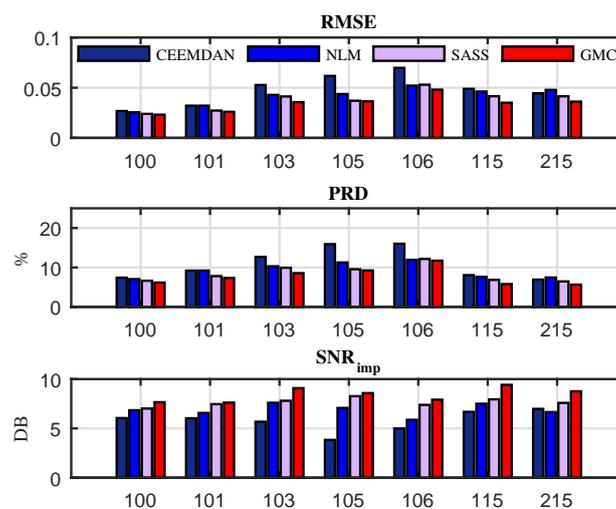

Figure 12. Performance evaluation over RMSE, PRD and SNR_{imp} criteria for different data sources.

To further show performance of the the proposed GMC-based method, we compared the SNR_{imp} percentage improvement of the GMC algorithm with the three referenced algorithms, i.e., SASS, NLM and CE-EMD-AN, under different noise levels. The percentage improvement is calculated according

to the following formula We designed an experiment to calculate the percentage of GMC increase in SNR_{imp} value relative to other algorithms. The calculation formula is as follows:

$$\text{Percentage improvement} = \frac{\text{SNR\_imp of GMC} - \text{SNR\_imp of the referenced algorithm}}{\text{SNR\_imp of the referenced algorithm}} \times 100\%. \quad (13)$$

The results are listed in Table 2. It shows that the performance of the proposed algorithm outperforms the referenced algorithms significantly.

Table 2. Percentage SNR_{imp} improvement relative to the three referenced algorithms.

	20 dB	15 dB	10 dB	5 dB	−5 dB	−10 dB
SASS	21.32	6.60	5.12	13.94	10.68	11.89
NLM	21.59	17.86	15.38	13.04	4.37	2.36
CE-EMD-AN	42.50	30.94	28.67	28.23	18.28	16.37

There is a question that if the proposed algorithm can be applied to the abnormal ECG signals, such as arrhythmia signal, and how it will perform. To answer this question, we apply the proposed algorithm to the normal heart beat signal and the arrhythmia signal and results are shown in Figures 13 and 14, respectively. As can be seen from Figure 13, all the four algorithms can remove noise effectively for the normal ECG signal. However, there is a huge difference in the type of signal that has an abnormal heart beat. As shown in Figure 14, after four seconds from the ECG signal (mitdb/230), the waveform begins to have turbulence and the T wave disappears significantly. After denoising, the output of the SASS algorithm is not smooth enough; the output of the CE-EMD-AN algorithm is seriously distorted in the period of 4 s–5 s; the output of the NLM algorithm has obvious sharp noise in the S-T segment during the 3 s–4 s period. Compared with these referenced algorithms, the proposed GMC algorithm achieves the best waveform. Therefore, the proposed GMC-based denoising algorithm can also be used with the arrhythmia signals.
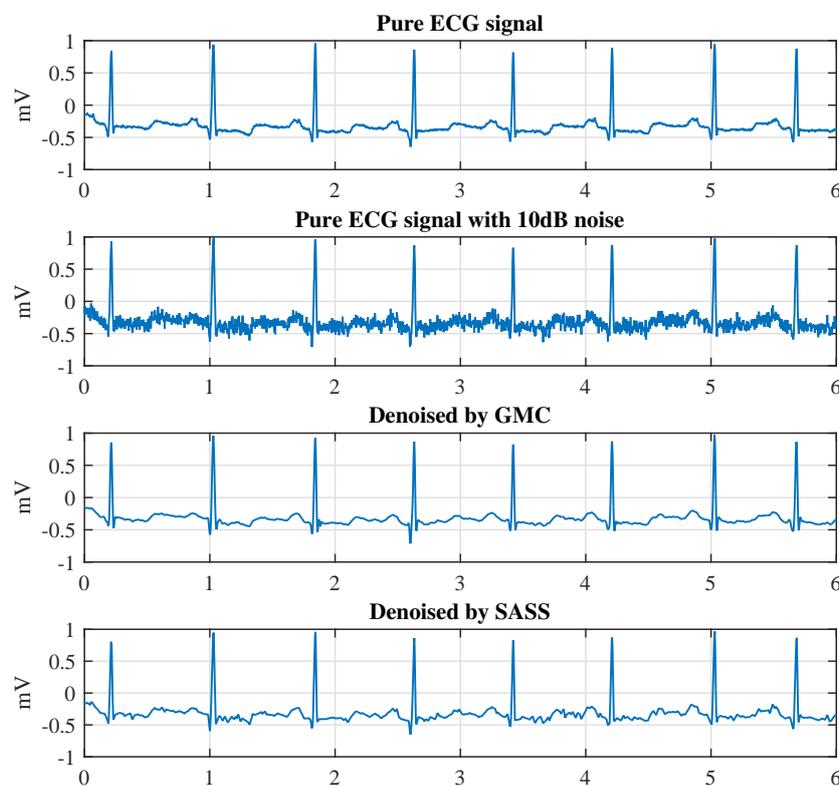

Figure 13. *Cont.*
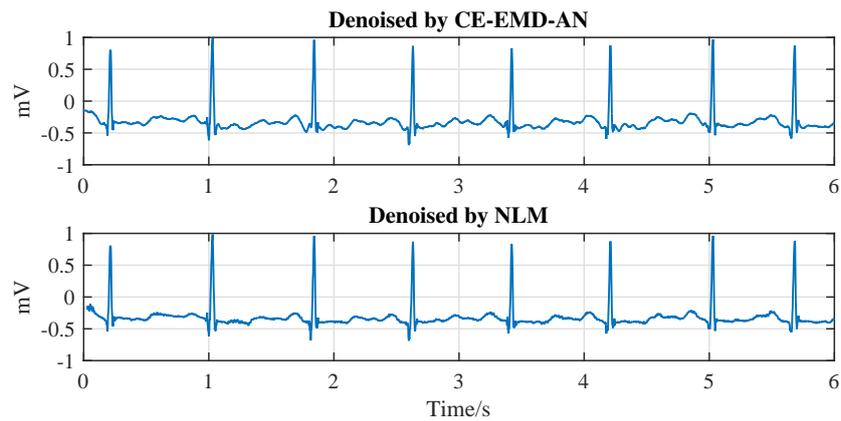
Figure 13. Comparison of waveform outputs of different algorithms for a normal ECG signal (No.100 in the MIT-BIH database).
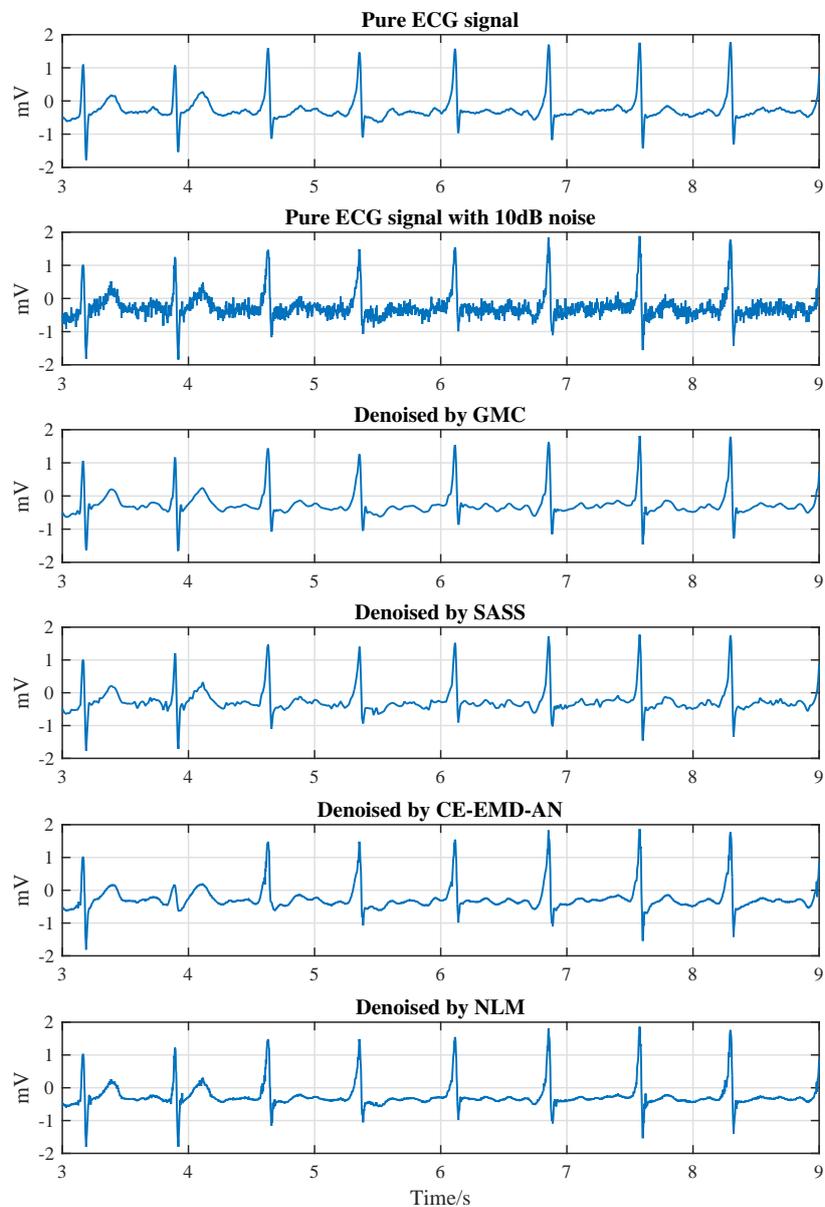

Figure 14. Comparison of waveform outputs of different algorithms for an arrhythmia ECG signal (No.230 in the MIT-BIH database).

6. Conclusions

We have investigated the ECG noise reduction techniques based on sparsity recovery in this paper. Through combining low-pass filtering and sparsity recovery, we proposed a novel sparse ECG denoising framework. The ℓ_1 penalty and the GMC penalty were utilized to construct sparsity recovery algorithms. Although the GMC penalty function was non-differentiable non-convex itself, it could be used to construct convex cost function and thus helps to obtain optimal solution. Since the GMC had the property of penalizing large values more than small values, but the degree of punishment is not as severe as ℓ_1-norm, it was utilized to overcome the drawback of underestimating the high-amplitude components for the ℓ_1-norm penalty. Experimental results demonstrated that the proposed GMC-based method overcame the underestimating problem of the ℓ_1-norm method and classical ECG denoising methods proposed in literature. Results showed that the GMC-based method outperforms the conventional algorithms such as the SASS algorithm and the EMD based method with respect to the SNR_{imp}, RMSE and PRD.

Abbreviations

The following abbreviations are used in this manuscript:

GMC	Generalized Minimax Concave Penalty
GSNC	Grey Spectral Noise Cancellation
GSNE	Grey Spectral Noise Estimation
ECG	Electrocardiogram
EMD	Empirical Mode Decomposition
SASS	Sparsity Assisted Signal Smoothing
NLM	Non-local Means
IMFs	Intrinsic Mode Functions
LTI	Linear Time Invariant
AWGN	Additive White Gaussian Noise
LPF	Low-pass Filter
HPF	High-pass Filter
STFT	Short Time Fourier Transform
BPD	Basis Pursuit Denoising
FBS	Forward-backward Spliting
SNR	signal-to-noise ratio
SNR_{imp}	signal-to-noise ratio improvement
RMSE	average of root mean square error
PRD	Percent root mean square difference

Appendix A. Filter Designing

In this paper, we take the LPFs to be zero-phase, non-causal, recursive filters considering that they filter relatively symmetric signals without introducing shifts in peak locations [29,38,49]. Recently, a novel procedure for defining such filters is given in [30]. We briefly sketch its key design idea here.

Given finite-length sequences p_n and q_n, two banded Toeplitz matrices P and Q are defined as

$$
P = \begin{bmatrix} p_1 & p_2 & p_3 & & & \\ & p_1 & p_2 & p_3 & & \\ & & & \ddots & & \\ & & & p_1 & p_2 & p_3 \end{bmatrix} ,
\tag{A1}
$$

$$
Q = \begin{bmatrix} q_1 & q_2 & q_3 & & & \\ & q_1 & q_2 & q_3 & & \\ & & & \ddots & & \\ & & & q_1 & q_2 & q_3 \end{bmatrix} .
\tag{A2}
$$

Based on these matrices, a linear filter matrix can be constructed as

$$
H = (Q^T Q + \alpha P^T P)^{-1} Q^T Q .
\tag{A3}
$$

The matrix H is approximately Toeplitz, and it represents approximately an LTI system with the frequency response

$$
H(e^{j\omega}) = \frac{|Q(e^{j\omega})|^2}{|Q(e^{j\omega})|^2 + \alpha |P(e^{j\omega})|^2} .
\tag{A4}
$$

$H(e^{j\omega})$ is zero-phase, and its transfer function is written as

$$
H(z) = \frac{Q(z)Q(z^{-1})}{Q(z)Q(z^{-1}) + \alpha P(z)P(z^{-1})} .
\tag{A5}
$$

After implementing Equation (A3) to the input data vector y (e.g., the noisy data), the vectorized output of the filter is written as

$$
x = Hy .
\tag{A6}
$$

An advantage of Equation (A6) is that transient effects at the start and end of the finite-length signal are intrinsically avoided. This means that the transient end-point artifacts is avoided.

In order to construct the transfer functions of the form Equation (A5), $P(z)$ and $Q(z)$ can be set as

$$
P(z) = (1 - z^{-1})^k ,
\tag{A7}
$$

$$
Q(z) = (1 + z^{-1})^k ,
\tag{A8}
$$

where k is the order of the filter. It is easy to verify that $P(1) = 0$ and $Q(-1) = 0$, $H(1) = 1$ and $H(-1) = 0$. Hence, the frequency response of the filter H has unity gain at $\omega = 0$ and zero at $\omega = \pi$.

Given Equations (A7) and (A8), the filter H is a discrete-time Butterworth filter with the frequency response

$$
H(e^{j\omega}) = \frac{\cos^{2k}(\omega/2)}{\cos^{2k}(\omega/2) + \alpha \sin^{2k}(\omega/2)} ,
\tag{A9}
$$

where α controls the cut-off frequency. For the 3 dB cutoff frequency, i.e., $H(e^{j\omega_c}) = 0.5$, α can be set as

$$
\alpha = 1/\tan^{2k}(w_c/2) .
\tag{A10}
$$

Appendix B. The STFT Used in the Sparsity Recovery

We use the STFT in the sparsity recovery block in Figure 1. Specifically, the input of the sparsity recovery block is cut as small data segments with length R. The segments are overlapped with a ratio of 50%. After padding R zeros behind the tail of the data segment, a STFT frame with length $N = 2R$ is constructed.

The constructed time-domain signal frame $d \in \mathcal{R}^N$ can be represented by the frequency-domain signal $x \in \mathcal{C}^N$ as

$$d = Ax \,, \tag{A11}$$

where $A \in \mathcal{C}^{N \times N}$ is the inverse discrete Fourier transform (IDFT) matrix defined as

$$A = \frac{1}{\sqrt{N}}[e^{j\frac{2\pi}{N}kn}]_{k=0,\ldots,N-1;n=0,\ldots,N-1} \,, \tag{A12}$$

The time-domain representation Equation (A11) can be transformed into the frequency-representation as

$$x = A^H d \,. \tag{A13}$$

It is noted specially that a window function with length R is applied to the segmented time-domain data, which is defined as

$$w(n) = \sin(\frac{n + 0.5}{R}\pi), n = 0,\ldots,R - 1 \,. \tag{A14}$$

After the frequency-to-time transform, this window should be removed by multiplexing its reciprocal.

References

1. Sayadi, O.; Shamsollahi, M.B. ECG denoising and compression using a modified extended Kalman filter structure. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 2240–2248. [CrossRef]
2. Van Alste, J.A.; Schilder, T. Removal of base-line wander and power-line interference from the ECG by an efficient FIR filter with a reduced number of taps. *IEEE Trans. Biomed. Eng.* **1985**, *BME-32*, 1052–1060. [CrossRef]
3. Camps, G.; Martinez, M.; Soria, E. Fetal ECG extraction using an FIR neural network. In Proceedings of the Computers in Cardiology 2001, Rotterdam, The Netherlands, 23–26 September 2001; pp. 249–252.
4. Belchandan, A.K.; Deshmukh, K.; Kumar, J. Removal of Noises in ECG Signal by using Digital FIR-IIR Filter in VHDL. *Digital Signal Process.* **2016**, *8*, 135–139.
5. Li, C.; Zheng, C.; Tai, C. Detection of ECG characteristic points using wavelet transforms. *IEEE Trans. Biomed. Eng.* **1995**, *42*, 21–28.
6. Addison, P.S. Wavelet transforms and the ECG: A review. *Physiol. Meas.* **2005**, *26*, R155. [CrossRef]
7. Singh, B.N.; Tiwari, A.K. Optimal selection of wavelet basis function applied to ECG signal denoising. *Digital Signal Process.* **2006**, *16*, 275–287.
8. Alfaouri, M.; Daqrouq, K. ECG signal denoising by wavelet transform thresholding. *Am. J. Appl. Sci.* **2008**, *5*, 276–281.
9. Vázquez, R.R.; Velez-Perez, H.; Ranta, R.; Dorr, V.L.; Maquin, D.; Maillard, L. Blind source separation, wavelet denoising and discriminant analysis for EEG artefacts and noise cancelling. *Biomed. Signal Process. Control* **2012**, *7*, 389–400.
10. Yadav, S.K.; Sinha, R.; Bora, P.K. Electrocardiogram signal denoising using non-local wavelet transform domain filtering. *IET Signal Proc.* **2015**, *9*, 88–96.
11. Goel, S.; Tomar, P.; Kaur, G. An optimal wavelet approach for ECG noise cancellation. *Int. J. Bio-Sci. Bio-Technol.* **2016**, *8*, 39–52.

12. Addison, P.S. *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*; CRC Press: Boca Raton, FL, USA, 2017.

13. Singh, O.; Sunkaria, R.K. ECG signal denoising via empirical wavelet transform. *Australas. Phys. Eng. Sci. Med.* **2017**, *40*, 219–229.

14. Kumar, A.; Komaragiri, R.; Kumar, M. A Wavelet Transform-Based Filter Bank Architecture for ECG Signal Denoising. In *Recent Findings in Intelligent Computing Techniques*; Springer: Berlin, Germany, 2018; pp. 249–255.

15. Blanco-Velasco, M.; Weng, B.; Barner, K.E. ECG signal denoising and baseline wander correction based on the empirical mode decomposition. *Comput. Biol. Med.* **2008**, *38*, 1–13. [CrossRef]

16. Lee, M.H.; Shyu, K.K.; Lee, P.L.; Huang, C.M.; Chiu, Y.J. Hardware implementation of EMD using DSP and FPGA for online signal processing. *IEEE Trans. Ind. Electron.* **2011**, *58*, 2473–2481. [CrossRef]

17. Kabir, M.A.; Shahnaz, C. Denoising of ECG signals based on noise reduction algorithms in EMD and wavelet domains. *Biomed. Signal Process. Control* **2012**, *7*, 481–489.

18. Jebaraj, J.; Arumugam, R. Ensemble empirical mode decomposition-based optimised power line interference removal algorithm for electrocardiogram signal. *IET Signal Proc.* **2016**, *10*, 583–591.

19. Kumar, S.; Panigrahy, D.; Sahu, P. Denoising of Electrocardiogram (ECG) signal by using empirical mode decomposition (EMD) with non-local mean (NLM) technique. *Biocybern. Biomed. Eng.* **2018**, *38*, 297–312. [CrossRef]

20. Rakshit, M.; Das, S. An efficient ECG denoising methodology using empirical mode decomposition and adaptive switching mean filter. *Biomed. Signal Process. Control* **2018**, *40*, 140–148.

21. Liu, S.H.; Hsieh, C.H.; Chen, W.; Tan, T.H. ECG Noise Cancellation Based on Grey Spectral Noise Estimation. *Sensors* **2019**, *19*, 798. [CrossRef]

22. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A* **1998**, *454*, 903–995. [CrossRef]

23. Kopsinis, Y.; McLaughlin, S. Development of EMD-based denoising methods inspired by wavelet thresholding. *IEEE Trans. Signal Process.* **2009**, *57*, 1351–1362.

24. Xiong, P.; Wang, H.; Liu, M.; Zhou, S.; Hou, Z.; Liu, X. ECG signal enhancement based on improved denoising auto-encoder. *Eng. Appl. Artif. Intell.* **2016**, *52*, 194–202. [CrossRef]

25. Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.* **2007**, *16*, 2080–2095. [CrossRef]

26. Ghafari, A.; Palangi, H.; Babaie-Zadeh, M.; Jutten, C. ECG denoising and compression by sparse 2D separable transform with overcomplete mixed dictionaries. In Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing (MLSP), Grenoble, France, 1–4 September 2009; pp. 1–6.

27. Ning, X.; Selesnick, I.W. ECG enhancement and QRS detection based on sparse derivatives. *Biomed. Signal Process. Control* **2013**, *8*, 713–723.

28. Zhou, Y.; Hu, X.; Tang, Z.; Ahn, A.C. Sparse representation-based ECG signal enhancement and QRS detection. *Physiol. Meas.* **2016**, *37*, 2093. [CrossRef]

29. Selesnick, I.W. Sparsity-assisted signal smoothing. In *Excursions in Harmonic Analysis*; Springer: Berlin, Germany, 2015; Volume 4, pp. 149–176.

30. Selesnick, I. Sparsity-assisted signal smoothing (revisited). In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 4546–4550.

31. Wang, S.; Selesnick, I.; Cai, G.; Feng, Y.; Sui, X.; Chen, X. Nonconvex Sparse Regularization and Convex Optimization for Bearing Fault Diagnosis. *IEEE Trans. Ind. Electron.* **2018**, *65*, 7332–7342. [CrossRef]

32. Candes, E.J.; Wakin, M.B.; Boyd, S.P. Enhancing sparsity by reweighted ℓ_1 minimization. *J. Fourier Anal. Appl.* **2008**, *14*, 877–905. [CrossRef]

33. Chouzenoux, E.; Jezierska, A.; Pesquet, J.C.; Talbot, H. A majorize-minimize subspace approach for ℓ_2-ℓ_0 image regularization. *SIAM J. Imaging Sci.* **2013**, *6*, 563–591. [CrossRef]

34. Soussen, C.; Idier, J.; Duan, J.; Brie, D. Homotopy based algorithms for ℓ_0-regularized least-squares. *IEEE Trans. Signal Process.* **2015**, *63*, 3301–3316. [CrossRef]

35. Selesnick, I. Sparse Regularization via Convex Analysis. *IEEE Trans. Signal Process.* **2017**, *65*, 4481–4494. [CrossRef]

36. Selesnick, I.; Farshchian, M. Sparse signal approximation via nonseparable regularization. *IEEE Trans. Signal Process.* **2017**, *65*, 2561–2575. [CrossRef]

37. Cai, G.; Selesnick, I.W.; Wang, S.; Dai, W.; Zhu, Z. Sparsity-enhanced signal decomposition via generalized minimax-concave penalty for gearbox fault diagnosis. *J. Sound Vib.* **2018**, *432*, 213–234.

38. Selesnick, I.W.; Graber, H.L.; Pfeil, D.S.; Barbour, R.L. Simultaneous Low-Pass Filtering and Total Variation Denoising. *IEEE Trans. Signal Process.* **2014**, *62*, 1109–1124.

39. Mathworks. *Maltab Help Documents: Zero-Phase Digital Filtering*; Mathworks: Natick, MA, USA, 2018.

40. Selesnick, I. L1-Norm Penalized Least Squares with SALSA. *Connexions* **2014**, 1–18. Available online: https://pdfs.semanticscholar.org/5a31/4bb73cea872da6ab27b680c30fa42ee5f37d.pdf (accessed on 10 April 2019).

41. Chen, S.; Saunders, M.A.; Donoho, D.L. Atomic decomposition by basis pursuit. *SIAM Rev.* **2001**, *43*, 129–159.

42. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Methodol.* **1996**, *58*, 267–288. [CrossRef]

43. Han, Z.; Li, H.; Yin, W. *Compressive Sensing for Wireless Networks*; Cambridge University Press: Cambridge, UK, 2013.

44. Benzi, M.; Golub, G.H.; Liesen, J. Numerical solution of saddle point problems. *Acta Numer.* **2005**, *14*, 1–137. [CrossRef]

45. Bauschke, H.H.; Combettes, P.L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*; Springer: New York, NY, USA, 2011; pp. 287–316.

46. Moody, G.B.; Mark, R.G. The impact of the MIT-BIH arrhythmia database. *IEEE Eng. Med. Biol. Mag.* **2002**, *20*, 45–50. [CrossRef]

47. Tracey, B.H.; Miller, E.L. Nonlocal means denoising of ECG signals. *IEEE Trans. Bio-Med. Eng.* **2012**, *59*, 2383–2386. [CrossRef] [PubMed]

48. Colominas, M.A.; Schlotthauer, G.; Torres, M.E. Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomed. Signal Process. Control* **2014**, *14*, 19–29. [CrossRef]

49. Ning, X.; Selesnick, I.W.; Duval, L. Chromatogram baseline estimation and denoising using sparsity (BEADS). *Chemom. Intell. Lab. Syst.* **2014**, *139*, 156–167. [CrossRef]