

Article

Not peer-reviewed version

A Denoising Algorithm Based on Spatial and Spectral Domain Filtering for Electrically Evoked Compound Action Potential Enhancement

[Fan-Jie Kung](#) *

Posted Date: 15 April 2025

doi: [10.20944/preprints202504.1267.v1](https://doi.org/10.20944/preprints202504.1267.v1)

Keywords: electrically evoked compound action potential (ECAP); panoramic ECAP (PECAP); log-spectral amplitude (LSA); root mean square error (RMSE)



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A Denoising Algorithm Based on Spatial and Spectral Domain Filtering for Electrically Evoked Compound Action Potential Enhancement

Fan-Jie Kung

Department of Electrical Engineering, National Taipei University of Technology; fjkung@mail.ntut.edu.tw

Abstract: The electrically evoked compound action potential (ECAP) is a physiological signal that is important for clinicians to assess a patient's hearing. A clean ECAP can be used to estimate the patient's auditory neural activity patterns and ECAP magnitude using the panoramic ECAP (PECAP) approach. However, noisy ECAPs can introduce significant errors in parameter estimation, especially at low signal-to-noise ratios (SNRs). To address this problem, a two-stage preprocessing denoising (TSPD) algorithm is developed. First, the ECAP matrix can be constructed using the forward-masking technique. Thus, the ECAP matrix can be viewed as an image for the first stage of noise reduction using the improved spatial median filter. Then, the denoised ECAP matrix can be expanded and rearranged into a vector for the second stage of noise reduction using the log-spectral amplitude (LSA) Wiener filter. Lastly, the enhanced ECAP vector is reconstructed to the ECAP matrix for aforementioned parameter estimation using PECAP. The results show that the proposed method, called PECAP-TSPD, improves the estimation of auditory neural activity patterns and ECAP magnitudes compared to those of the baseline (PECAP) and unprocessed ECAP signals in terms of the normalized root mean square error (RMSE).

Keywords: electrically evoked compound action potential (ECAP); panoramic ECAP (PECAP); log-spectral amplitude (LSA); root mean square error (RMSE)

1. Introduction

The electrically evoked compound action potential (ECAP) is the combined response of auditory nerve fibers to electrical stimulation by a cochlear implant (CI) [1,2]. The ECAP is a critical signal for clinicians to evaluate the functionality of a patient's auditory nerve fibers after CI surgery [3,4]. In the absence of feedback from CI users, the ECAP model can be of great value to clinicians assessing the hearing performance of CI users [5,6]. ECAP amplitude can be computerized and estimated using temporal or spatial methods [7–9]. The ECAP matrix can be constructed by measuring the ECAPs generated by the different electrode positions of the masker and probe stimuli. Such a method is referred to as the forward-masking technique [10]. The panoramic ECAP (PECAP) method [11,12] in the literature has shown that using a clean ECAP matrix can approximate auditory activity patterns and ECAP magnitudes [11,12], further assisting clinicians in evaluating a patient's speech perception. Nevertheless, the use of noisy ECAP matrices, particularly when the signal-to-noise ratio (SNR) is less than 10 dB, has the potential to introduce inaccuracies in the estimation of auditory activity patterns and ECAP magnitudes [11,12].

The aforementioned reason necessitates the implementation of noise reduction techniques on ECAP matrices prior to estimating the auditory activity patterns and ECAP magnitudes. These techniques can be classified into three primary categories: spatial filtering, temporal filtering, and spectral filtering. In the context of image signal processing, the mean and median filters represent well-established examples of spatial filtering [13–15]. The mean filter is a linear estimator that is employed to mitigate the adverse effects of noise in images by eliminating random fluctuations. The principal disadvantage of the mean filter is that it results in image blurring, particularly at low SNRs.



The median filter, which demonstrates superior noise removal performance in comparison to the mean filter, is a nonlinear estimator [16]. The median filter's masking size tends to increase with elevated noise levels, which can result in the loss of information from the original images. Thus, an improved median filtering algorithm (I-median) [17] is developed that combines the advantages of mean and median filters with the objective of achieving superior denoising in adverse environments. Specifically, if the current pixel is less than the average of the mask, the pixel is replaced with the median of the mask. Otherwise, the pixel keeps its original value [17].

In the field of time-domain signal analysis, adaptive filtering (AF) can be used to minimize noise. The AF method primarily uses algorithms such as least mean square (LMS) [18] and recursive least square (RLS) [19] to iteratively adjust the weights, thereby reducing the adverse effects of noise on the signal. The advantage of AF methods is its near real-time implementation [20]. However, a critical aspect of this method lies in the reference signal source, as clean signals, such as speech, are almost nonexistent in certain real-world scenarios. In addition, the choice of step size and filter order, which directly affect the rate of error convergence, is an important issue to address [21]. Finding the right balance is essential for noise reduction when AF is used.

Wiener filtering is an effective noise reduction technique in spectral domain signal analysis [22,23]. It operates by estimating the power spectral density (PSD) of the noise to enhance the signal. However, at low SNRs, the accuracy of the estimated noise PSD may decline, directly impacting noise reduction performance. To address this issue, a Wiener-based noise reduction algorithm known as log-spectral amplitude (LSA) Wiener filtering [24–26], has been developed to minimize the mean square error of the logarithmic spectrum. This approach helps alleviate the musical noise and improve the speech quality under low SNR conditions.

Recently, deep learning techniques have also been applied in the field of speech enhancement [27–30]. One such technique is the end-to-end convolutional recurrent neural network (CRN) [31], which enables near real-time implementation using a single microphone. This CRN algorithm can suppress noise under adverse conditions, such as a -5 dB SNR. However, the main drawback of these learning-based approaches is the substantial amount of data acquisition and preprocessing required [27,28,32], which can be both time- and resource- intensive. Additionally, the noise reduction performance may degrade when encountering unseen scenarios.

Given the compromised estimation of activation patterns and ECAP magnitudes using PECAP under low SNR conditions, this work proposes a method that combines PECAP with a two-stage preprocessing denoising algorithm (TSPD), referred to as PECAP-TSPD. Specifically, the ECAP signals are first denoised by TSPD and then used to estimate the neural parameters by PECAP. In the first stage of TSPD, the I-median algorithm is used to reduce the random noise of an ECAP matrix, since the ECAP matrix can be considered as an image. In the second step of TSPD, the denoised ECAP matrix is expanded and rearranged as a $1 \times NN$ vector, where N is the total number of electrodes. The rearranged vector can then be used as the input of the one-dimensional signal to LSA Wiener filtering for residual noise reduction. The PECAP-TSPD algorithm improves the accuracy of the estimates for key parameter, e.g., neural health and current spread. As a result, the refined approach leads to more reliable estimates of neural activation patterns and ECAP magnitudes, improving the effectiveness of neural assessments by clinicians. The normalized root mean square error (RMSE) serves as an objective measure in this work to evaluate the performance of the unprocessed data, PECAP-processed data, and PECAP-TSPD-processed data under various SNR conditions, densities, and scenarios. It also serves to highlight the accuracy of estimating ECAP magnitudes and activation patterns.

2. Panoramic ECAP Method

The panoramic ECAP (PECAP) method consists of two procedures: the forward-masking method and neural parameter estimation using sequential quadratic programming (SQP) [10–12]. The forward-masking method uses a probe and a masker to stimulate auditory neurons, generating the ECAP signal and reducing artifact. This process constructs a matrix in which each element

represents the measured ECAP signal from each pair of probe and masker positions, as illustrated in Figure 1. In Figure 1, the x-axis represents the position of the masker, and the y-axis represents the position of the probe. Figure 1 shows the simulated ECAP signal, illustrating the inverse relationship between the amplitude of the ECAP and the distance from the probe to the masker. When the positions of the probe and the masker are closer, the amplitude of the ECAP signal is larger. Conversely, when the probe and masker are farther apart, the amplitude of the ECAP signal is smaller. This phenomenon can be explained in Figure 2 [6]. In Figure 2, the neural activity pattern is assumed to follow a Gaussian distribution. The ECAP signal can be regarded as the overlapping area between two Gaussian distributions, corresponding to the neural activity patterns stimulated by the probe and masker, respectively. The overlap is maximized when the positions of the probe and the masker coincide. In contrast, the overlap becomes minimal when the positions of the probe and the masker are distant from each other.

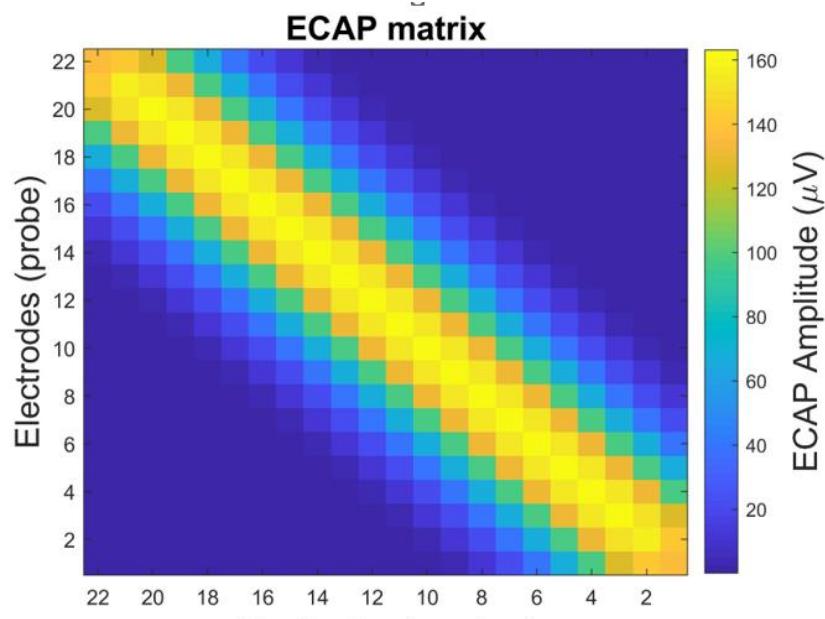


Figure 1. An illustration of an ECAP matrix simulation. The x-axis represents the position of the masker (from Electrode 22 to Electrode 1), and the y-axis represents the position of the probe (from Electrode 1 to Electrode 22) [12].

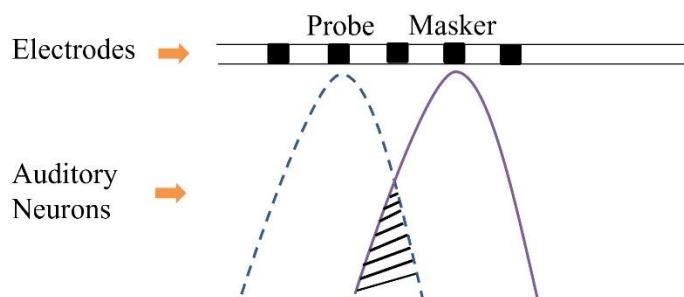


Figure 2. Schematic illustrating the overlapping area stimulated by the probe and the masker. Gaussian distributions represent the auditory neuron response, and the shaded overlapping indicates the ECAP response [6].

The auditory neural activity pattern is assumed to be a Gaussian distribution as follows:

$$a_i(k) = \alpha_i \eta_i \exp \left\{ -\frac{(k - \mu_i)^2}{2\sigma_i^2} \right\} \quad (1)$$

where $i = 1, 2, \dots, N$ denotes the i th electrode. N denotes the total number of electrodes. k is the position along the cochlea. Ideally, k is a range from $-\infty$ to $+\infty$. In this work, the range of k is from 1 to K in which K is set to N , the total number of electrode. α_i and η_i are the auditory neuron amplitude and neural health of the i th electrode, respectively. $\exp\{\cdot\}$ is an exponential function. u_i and σ_i represent the mean and standard deviation (also referred to as current spread¹⁰ in physiological signals) of the Gaussian pattern for the i th electrode, respectively. In this work, $u_i = i$ and α_i is a prior information. Under an assumption that the ECAP signal is the overlap between the two auditory neuron responses of the probe and masker stimuli, the ECAP signal can be formulated as

$$M(p, m) = \sum_{k=1}^K a_p(k)a_m(k), \quad (2)$$

where $p = 1, 2, \dots, N$ is the index indicating that the p th electrode is stimulated by the probe. $m = 1, 2, \dots, N$ is the index indicating that the m th electrode is stimulated by the masker. Equation (2) can also be expressed as a matrix form [12].

$$\mathbf{M} = \sqrt{\mathbf{A}\mathbf{A}^T}, \quad (3)$$

in which $\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_N]^T \in \Re^{N \times K}$ is an auditory neuron pattern matrix for each auditory neuron vector $\mathbf{a}_i = [a_i(1) \ a_i(2) \ \cdots \ a_i(K)]^T \in \Re^K$. Because \mathbf{A} is a symmetric matrix, the matrix of the ECAP signal in Equation (3) is also a symmetric matrix. The part of random noise can be alleviated using the following equation:

$$\mathbf{M}_o = \frac{\mathbf{M} + \mathbf{M}^T}{2}, \quad (4)$$

where \mathbf{M}_o represents the denoised matrix of the ECAP signal. To further estimate η_i and σ_i , the SQP, a constrained optimization algorithm, is utilized. Computerization of the parameter estimation of the ECAP matrix is illustrated in Figure 3.

In Figure 3, ε_M , which denotes the normalized root mean square error between the measured and estimated ECAP matrices, is formulated as

$$\varepsilon_M = \frac{\sqrt{\frac{1}{N^2} \left| M_s(p, m) - \hat{M}_s(p, m) \right|^2}}{\bar{M}_s}, \quad (5)$$

where \bar{M}_s represents the maximum absolute value of \mathbf{M}_s . Different SNR and density conditions are also introduced to assess the parameter estimation capability using SQP in this work.

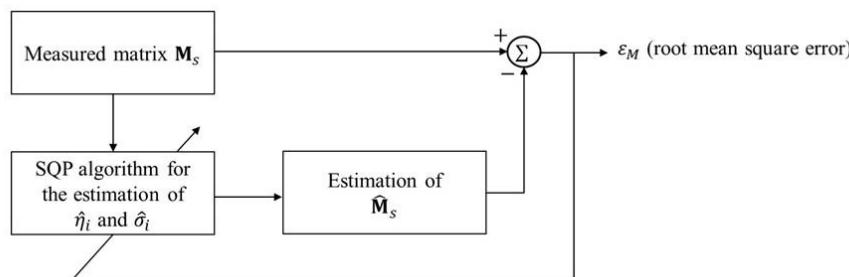


Figure 3. The block diagram of neural parameter estimation in the ECAP matrix using the SQP algorithm [6,12].

Figure 4 depicts a structure designed to test the noise resistance performance using the SQP algorithm. The \mathbf{M}_s in Figures 3 and 4 can be replaced by \mathbf{M}_o , as described in Equation (4) for random noise reduction. In addition, the normalized root mean square error between the ground truth \mathbf{A} and the estimated $\hat{\mathbf{A}}$ can be calculated, as shown in Equation (6).

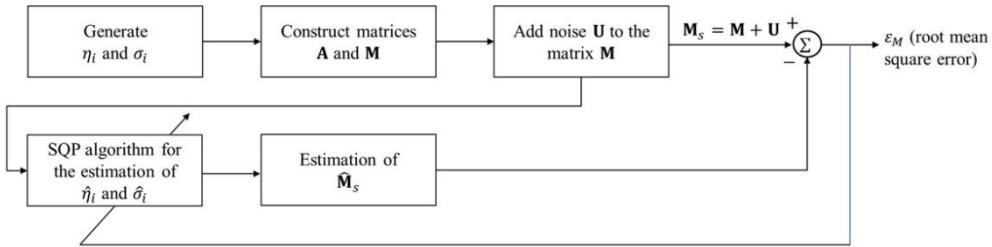


Figure 4. The block diagram of parameter estimation with additive noise scenarios using the SQP algorithm [11,12].

$$\varepsilon_A = \sqrt{\frac{1}{N^2} |A(p, m) - \hat{A}(p, m)|^2 / \bar{A}} \quad (6)$$

where \bar{A} denotes the maximum absolute value of A . After estimating the parameters of η_i and σ_i , the denoised ECAP matrix can be reconstructed using Equation (3). The above procedures can eliminate most part of the noise. However, in low SNR scenarios, the distortion increases due to the error estimation of η_i and σ_i . Therefore, the preprocessing algorithm is developed below.

3. Proposed Method

3.1. First Stage Noise Reduction Processing

In light of the detrimental effect of noise on the estimation of neural parameters, this work proposes a two-stage preprocessing denoising (TSPD) algorithm for the ECAP matrix, prior to the PECAP algorithm. First, the ECAP matrix is treated as an image. In the first stage of TSPD, the improved median filtering (I-median) algorithm is applied to reduce noise, as shown in Equation (7).

$$M_I(p, m) = \begin{cases} o(p, m) & M(p, m) < \mu(p, m) \\ M(p, m) & \text{otherwise} \end{cases}, \quad (7)$$

where $o(p, m)$ and $\mu(p, m)$ denote the processed values using the median and mean filters at the positions of (p, m) , respectively. $M_I(p, m)$ denotes the processed result using I-median. The masker sizes of median and mean filtering are set to 3×3 in this work. Equation (7) describes that if the ECAP value at the positions (p, m) is less than the mean filter processing value, the ECAP value is considered noise and can be replaced by the median filter processing value. Conversely, if the ECAP value is greater, it is retained. The processed ECAP matrix is expressed as

$$\mathbf{M}_I = \begin{bmatrix} M_I(1,1) & \cdots & M_I(1,N) \\ \vdots & \ddots & \vdots \\ M_I(N,1) & \cdots & M_I(N,N) \end{bmatrix}, \quad (8)$$

The I-Median algorithm is suitable for removing low to medium density noise from an image. In some cases, however, the noise is distributed across all pixels of an image. To further deal with this type of noise, and assuming that the ECAP signals are the overlaps between the probe and masker stimuli. This work expands \mathbf{M}_I into a $1 \times NN$ vector in accordance with the following rule:

1. Calculate the absolute value of index p minus index m .

$$\mathbf{B} = \begin{bmatrix} |1 - 1| & \cdots & |1 - N| \\ \vdots & \ddots & \vdots \\ |N - 1| & \cdots & |N - N| \end{bmatrix} = \begin{bmatrix} 0 & \cdots & |1 - N| \\ \vdots & \ddots & \vdots \\ |N - 1| & \cdots & 0 \end{bmatrix}, \quad (9)$$

where \mathbf{B} is an index matrix that records the absolute value of position difference between p and m .

2. record $|p - m|$ for each element in \mathbf{B} and concatenate each row into a long vector.

$$\mathbf{b} = [b(1,1), \dots, b(1,N), b(2,1), \dots, b(2,N), \dots, b(N,1), \dots, b(N,N)] \in \Re^{1 \times NN}, \quad (10)$$

where $\mathbf{b}(\mathbf{p}, \mathbf{m}) = |\mathbf{p} - \mathbf{m}|$.

3. Conduct a descending order of \mathbf{b} and record the descending order index.

$$(\mathbf{i}_d, \bar{\mathbf{b}}) = \text{descend}(\mathbf{b}), \quad (11)$$

where $\text{descend}(\cdot)$ represents the descending order operation. $\bar{\mathbf{b}}$ is vector based on the descending order results of \mathbf{b} . \mathbf{i}_d is the index vector corresponding to $\bar{\mathbf{b}}$.

4. The desired vector then can be obtained using the following equation:

$$\mathbf{m}_I = [M_I(\mathbf{i}_d(1)), M_I(\mathbf{i}_d(2)), \dots, M_I(\mathbf{i}_d(NN))] \in \Re^{1 \times NN}, \quad (12)$$

By examining the ECAP signal in Figures 1 and 2, the maximum can be reached when the probe and masker positions are close, and the minimum can be reached when the probe and masker positions are far apart. Therefore, in the previous steps, the noise (where ECAP signal is weak) would be placed in the first part of the vector. The target signal (where strong ECAP signal) can then be placed in the last part of the vector. The reordering processing is part of the second stage of TSPD, which employs log-spectral amplitude (LSA) Wiener filtering for improved noise reduction.

3.2. Second Stage Noise Reduction Processing

In the second stage of TSPD, LSA Wiener filtering is utilized to address the residual noise in the vector \mathbf{m}_I , which is treated as a discrete-time series input. The first step is to transform \mathbf{m}_I into a short-time Fourier transform (STFT) domain, as shown in Equation (13).

$$Y_m(l, k) = \sum_{n=0}^{N_{FFT}-1} M_I(n + lN_{hop}) w(n) W_{N_{FFT}}^{kn}, \quad (13)$$

where $Y_m(l, k)$ is the STFT signal of \mathbf{m}_I . l and k denote the time frame and frequency indices, respectively. N_{FFT} is the size of fast Fourier transform (FFT), N_{hop} is the hop size, and $w(n)$ is the window for short-time signal analysis. $W_{N_{FFT}}^k = \exp\{-j 2\pi k/N_{FFT}\}$ is the N_{FFT} th root of unity with $k = 0, 1, \dots, N_{FFT} - 1$. The LSA Wiener filter aims at minimizing the log-spectral amplitude

$$J = \arg \min_{H(l,k)} E \left\{ \left| \log_e |X(l, k)| - \log_e |\hat{X}(l, k)| \right|^2 \right\}, \quad (14)$$

where J is the cost function to minimize the mean square error of log-spectral amplitudes $\log_e |X(l, k)|$ and $\log_e |\hat{X}(l, k)|$. $X(l, k)$ and $\hat{X}(l, k)$ are the clean signal and estimated clean signal spectrums, respectively. The optimal solution is

$$H_{LSA}(l, k) = \frac{\xi(l, k)}{1 + \xi(l, k)} \exp \left\{ \frac{1}{2} \int_{v(l,k)}^{\infty} \frac{e^{-t}}{t} dt \right\}, \quad (15)$$

where $\xi(l, k) = E\{|X(l, k)|^2\}/P_u(l, k)$ is the prior SNR with $P_x(l, k) = E\{|X(l, k)|^2\}$ being a clean signal PSD and $P_u(l, k) = E\{|U_r(l, k)|^2\}$ being a residual noise PSD. $v(l, k) = (\xi(l, k)/1 + \xi(l, k))\gamma(l, k)$ with $\gamma(l, k) = |Y_m(l, k)|^2/P_u(l, k)$ representing the posterior SNR. $\xi(l, k)$ can be estimated using the decision-directed approach as described below:

$$\hat{\xi}(l, k) = \alpha \frac{|X(l-1, k)|^2}{P_u(l-1, k)} + (1 - \alpha) \max\{\gamma(l, k) - 1, \varepsilon_0\}, \quad (16)$$

where α is a forgetting factor. $P_u(l, k)$ can be estimated and updated using the following log-likelihood ratio criterion:

$$\Lambda(l, k) = \ln \frac{f(Y_m(l, k)|H_1)}{f(Y_m(l, k)|H_0)} = -\ln(1 + \xi(l, k)) + v(l, k), \quad (17)$$

where

$$f(Y_m(l, k)|H_1) = \frac{1}{\pi(P_x(l, k) + P_u(l, k))} \exp\left\{-\frac{|Y_m(l, k)|^2}{P_x(l, k) + P_u(l, k)}\right\} \quad (18)$$

is the conditional probability density function (PDF) of $Y_m(l, k)$ given that the event (H_1) of the ECAP signal occurs.

$$f(Y_m(l, k)|H_0) = \frac{1}{\pi P_u(l, k)} \exp\left\{-\frac{|Y_m(l, k)|^2}{P_u(l, k)}\right\} \quad (19)$$

is the conditional PDF of $Y_m(l, k)$ assuming only noise occurs at event H_0 . $v(l, k) = \gamma(l, k)\xi(l, k)/(1 + \xi(l, k))$. If $\sum_{l=1}^{N_t} \Lambda(l, k)$ is less than a small value ε_1 , the noise PSD can be updated using the following recursive averaging.

$$\hat{P}_u(l, k) = \alpha \hat{P}_u(l-1, k) + (1 - \alpha)|Y_m(l, k)|^2. \quad (20)$$

The estimated clean ECAP signal can be obtained using the following equation:

$$\hat{X}(l, k) = H_{LSA}(l, k)Y_m(l, k). \quad (21)$$

The complete LAS Wiener filtering procedures used in the second noise reduction stage are listed in Table 1.

In Table 1, N_u is the number of time frames used to represent the presence of noise. Next, the processed row vector $\hat{X}(l, k)$ is converted into using the inverse STFT the time domain signal $\hat{x}(n)$ which can be reconstructed into the matrix format, as described below:

$$\tilde{\mathbf{M}}_I = \begin{bmatrix} \hat{x}(n'|\mathbf{i}_d(n') = 1) & \cdots & \hat{x}(n'|\mathbf{i}_d(n') = N) \\ \hat{x}(n'|\mathbf{i}_d(n') = N+1) & \ddots & \hat{x}(n'|\mathbf{i}_d(n') = 2N) \\ \vdots & & \vdots \\ \hat{x}(n'|\mathbf{i}_d(n') = (N-1)N+1) & \cdots & \hat{x}(n'|\mathbf{i}_d(n') = NN) \end{bmatrix}. \quad (22)$$

where $\mathbf{i}_d(n')$ is defined in Equation (11) with n' being a range of 1 to NN . In this work, the mean filter is applied in $\tilde{\mathbf{M}}_I$ if the SNR value is below 4 dB, in order to leverage the advantage of random noise removal at low SNRs [15].

Table 1. Steps involved in the LSA Wiener filtering algorithm.

Step 1. Initialization of $P_u(k) = \left| \frac{1}{N_u} \sum_{l=1}^{N_u} |Y_m(l, k)| \right|^2$ for each frequency bin

For each l and k :

Step 2. Estimation of $\gamma(l, k)$

If $l = 1$, then $\gamma(l, k) = \frac{|Y_m(l, k)|^2}{P_u(k)}$, else $\gamma(l, k) = \frac{|Y_m(l, k)|^2}{P_u(l-1, k)}$

Step 3. Estimation of $\hat{\xi}(l, k)$ using Equation (16)

If $l = 1$, then Equation (16) can be rewritten as $\hat{\xi}(l, k) = (1 - \alpha)\max\{\gamma(l, k) - 1, \varepsilon_0\}$

Step 4. check the VAD criterion

If $\sum_{l=1}^{N_t} \Lambda(l, k) < \varepsilon_1$, then using Equation (20) for updating $P_u(l, k)$

Step 5. Calculation of $H_{LSA}(l, k)$ using Equation (15)

Step 6. Calculation of $\hat{X}(l, k)$ using Equation (21)

End for

4. Settings and Results

Two types of noise - random noise and impulse noise - are used to evaluate the performance of the proposed PECAP-TSPD algorithm. Twelve SNR levels, -5 dB, -2 dB, 1 dB, 4 dB, 7 dB, 10 dB, 13 dB, 16 dB, 19 dB, 22 dB, 25 dB, and 100 dB (representing the clean ECAP signal situation for random noise case) are used in the random noise. Four densities, 10%, 20%, 30%, and 40%, are used in the impulse noise. Normalized RMSE is used as an objective quality measure to calculate the error between the ground truth and the estimated results (neural health and ECAP amplitude). Seven different combinations of neural health and current spread are listed in Table 2. The results of the ECAP matrices before and after processing at -5 dB of SNR are shown in Figure 5.

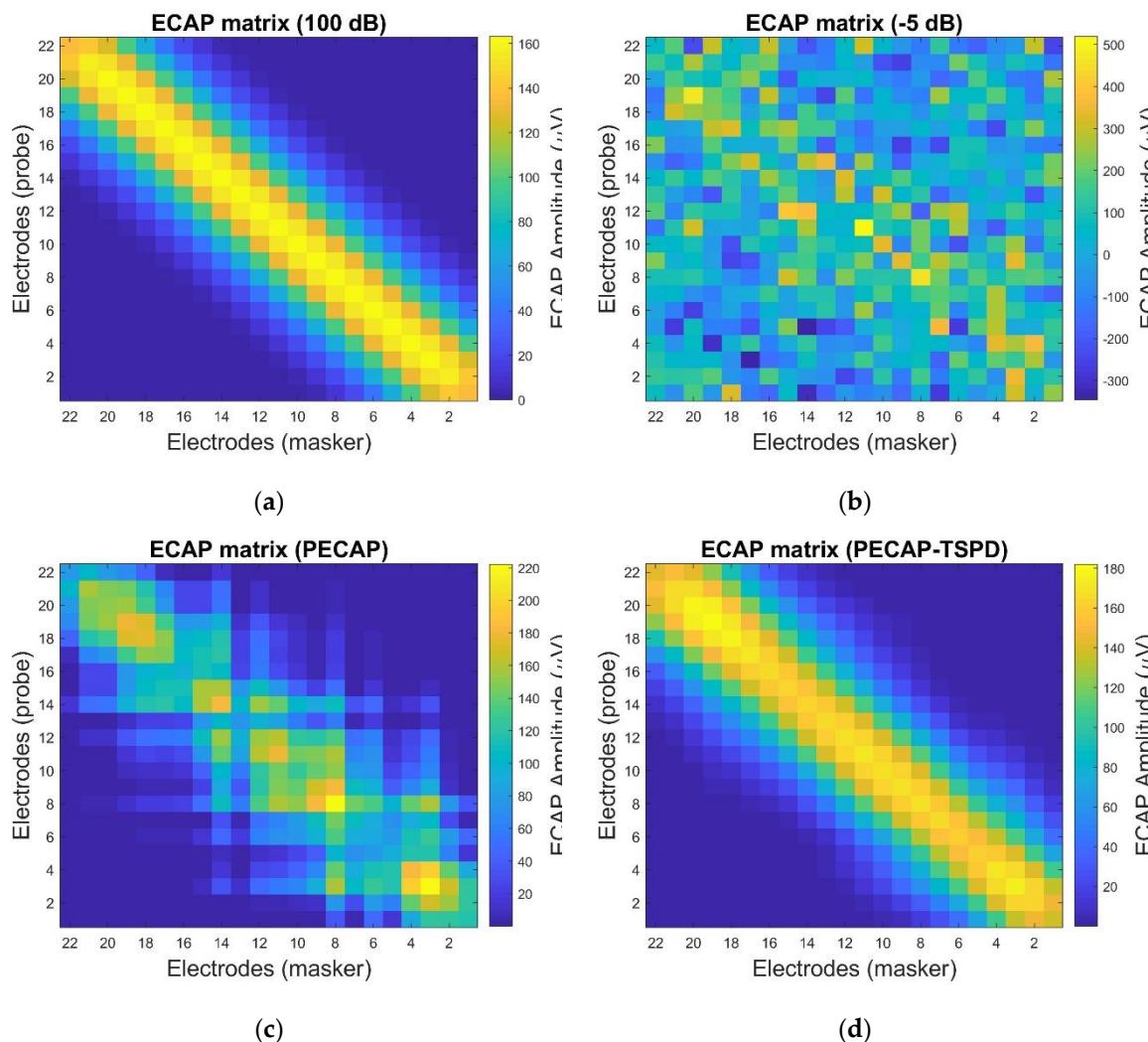


Figure 5. The results of (a) ECAP matrix at 100 dB SNR, (b) noisy ECAP matrix at -5 dB SNR, (c) processed ECAP matrix using the PECAP method at -5 dB SNR, and (d) processed ECAP matrix using the TSPD method at -5 dB. The neural health and current spread settings are scenario 1 listed in Table 2.

In Figure 5(b), the ECAP matrix is filled with random noise, thereby, making it difficult to observe the pristine measured ECAP data, as depicted in Figure 5(a). The PECAP and PECAP-TSPD methods can mitigate the detrimental effects of extremely noisy environments to recover the clean ECAP matrix shown in Figures 5(c) and (d). The images of Figure 5(a) and (d) are almost identical, further presenting the satisfactory performance of PECAP-TSPD. The results of the normalized RMSE of the ECAP magnitude (ε_M) and neural activity pattern (ε_A) are depicted in Figure 6.

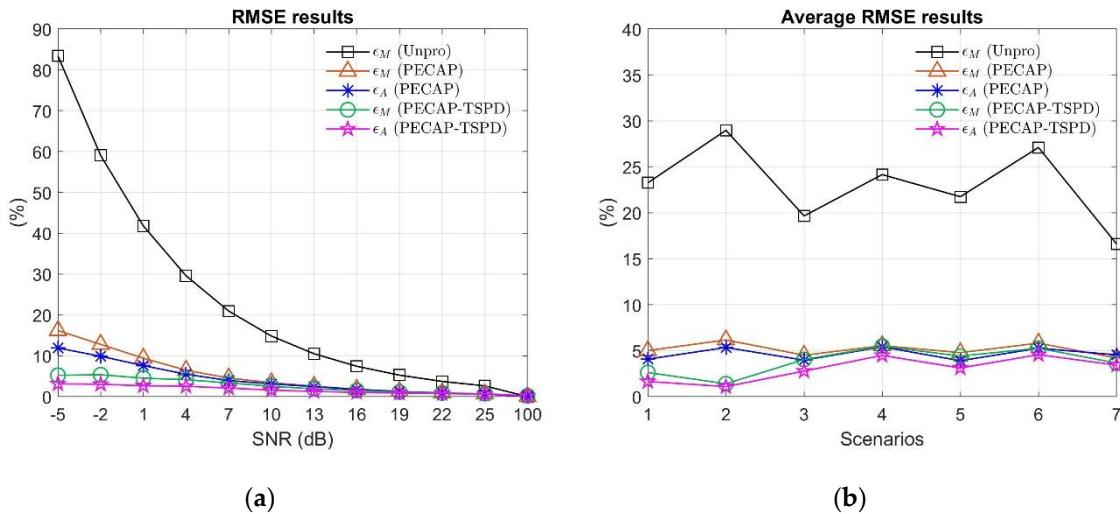


Figure 6. The normalized RMSE and average normalized RMSE results of ECAP magnitudes and neural activity patterns under twelve SNR conditions (a) in scenario 1 and (b) in scenarios 1 to 7, respectively. The unprocessed ECAP matrices are used for comparison with the baseline PECAP and proposed PECAP-TSPD methods. The parameter settings for each scenario are shown in Table 2.

Table 2. Different settings of neural health and current spread.

Scenario 1: $\eta_i = 1$, $\sigma_i = 1.5$, $i = 1, 2, \dots, N$, where $N = 22$ in this study.

Scenario 2: $\eta_i = 1$, $\sigma_i = 2.5$, $i = 1, 2, \dots, N$.

Scenario 3: $\eta_{i'} = 1$, $i' = 1, 2, \dots, 13, 21, 22$.

$\eta_{14} = \eta_{20} = 0.75$, $\eta_{15} = \eta_{19} = 0.50$, $\eta_{16} = \eta_{18} = 0.25$, $\eta_{17} = 0.10$.

$\sigma_i = 1.5$, $i = 1, 2, \dots, N$.

Scenario 4: $\eta_{i'} = 1$, $i' = 1, 2, \dots, 13, 21, 22$.

$\eta_{14} = \eta_{20} = 0.75$, $\eta_{15} = \eta_{19} = 0.50$, $\eta_{16} = \eta_{18} = 0.25$, $\eta_{17} = 0.10$.

$\sigma_i = 2.5$, $i = 1, 2, \dots, N$.

Scenario 5: $\eta_{i'} = 1$, $i' = 1, 2, \dots, 18$.

$\eta_{19} = 0.75$, $\eta_{20} = 0.50$, $\eta_{21} = 0.25$, $\eta_{22} = 0.10$.

$\sigma_i = 1.5$, $i = 1, 2, \dots, N$.

Scenario 6: $\eta_{i'} = 1$, $i' = 1, 2, \dots, 18$.

$\eta_{19} = 0.75$, $\eta_{20} = 0.50$, $\eta_{21} = 0.25$, $\eta_{22} = 0.10$.

$\sigma_i = 2.5$, $i = 1, 2, \dots, N$.

Scenario 7: $\eta_{i'} = 0.5$, $i' = 1, 2, \dots, 12, 15, 22$.

$\eta_{13} = 0.6$, $\eta_{14} = 0.7$, $\eta_{16} = \eta_{21} = 0.4$, $\eta_{17} = \eta_{20} = 0.3$, $\eta_{18} = \eta_{19} = 0.2$.

$\sigma_1 = 1.5$, $\sigma_{i''} = 2.5 - 0.05(i'' - 1)$, $i'' = 2, \dots, N$

Figure 6(a) shows the normalized RMSE of the unprocessed ECAP signals, ECAP signals processed by PECAP and PECAP-TSPD algorithms for different SNRs in Scenario 1. The normalized RMSE of the magnitude of the unprocessed ECAP signals at -5 dB SNR increases to 83.39%, which is comparably higher than those of the processed ECAP signals by PECAP (16.17%) and PECAP-TSPD (5.23%), indicating the need for ECAP signal processing. When comparing the ϵ_A between PECAP

and PECAP-TSPD processing ECAP signals, the values of ε_A by PECAP-TSPD are all smaller than those by PECAP except the 100 dB SNR case, where ε_A are 0.0043% and 0.1885% for PECAP and PECAP-TSPD, respectively. Figure 6(b) presents the curve of the average normalized RMSE, which is formulated as follows for each scenario:

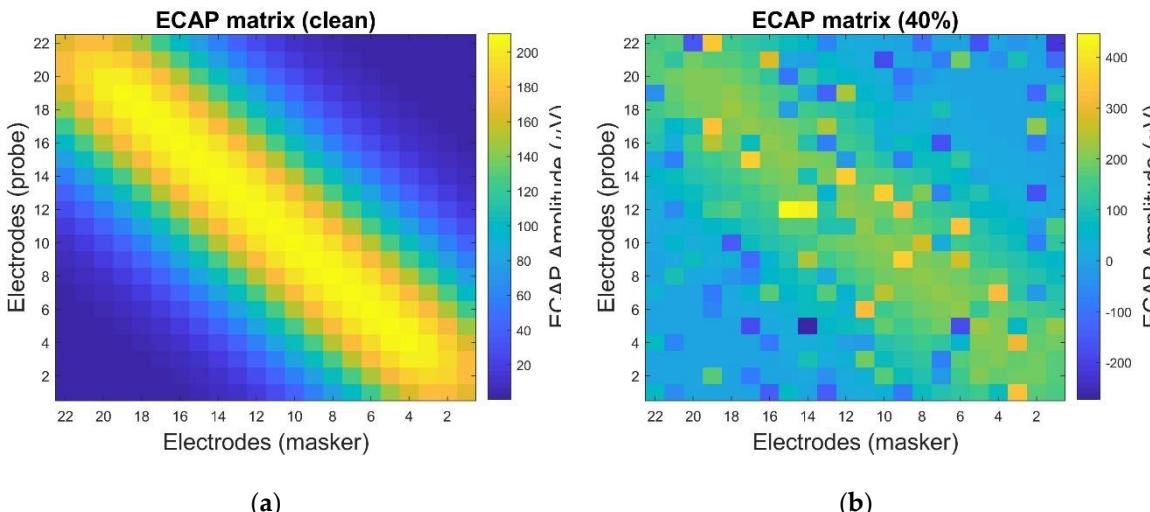
$$\bar{\varepsilon}_M = \sum_{k=1}^{N_{SNR}} \frac{\sqrt{\frac{1}{N^2} |M_{s,k}(p, m) - \hat{M}_{s,k}(p, m)|^2}}{\bar{M}_{s,k}}, \quad (23)$$

where $N_{SNR} = 12$ represents the total number of SNR conditions in this work. $M_{s,k}(p, m)$ and $\hat{M}_{s,k}(p, m)$ represent each index (p, m) value of each SNR condition (k) in the pristine and estimated ECAP matrices, respectively. $\bar{M}_{s,k}$ is the maximum absolute value of $\mathbf{M}_{s,k}$ which denotes the pristine ECAP matrix of each SNR condition. Similar to $\bar{\varepsilon}_M$, $\bar{\varepsilon}_A$ is described as

$$\bar{\varepsilon}_A = \sum_{k=1}^{N_{SNR}} \sqrt{\frac{1}{N^2} |A_{s,k}(p, m) - \hat{A}_{s,k}(p, m)|^2 / \bar{A}_{s,k}}, \quad (24)$$

in which $A_{s,k}(p, m)$ and $\hat{A}_{s,k}(p, m)$ denote each index value of each SNR condition in the pristine and estimated neural activity patterns, respectively. $\bar{A}_{s,k}$ is the maximum absolute value of $\mathbf{A}_{s,k}$ which denotes the pristine neural activity pattern of each SNR condition. Compared to the unprocessed and processed ECAP signals, the maximum values of $\bar{\varepsilon}_M$ are 6.17% and 5.48% for PECAP and PECAP-TSPD, respectively. In contrast, the maximum value of the unprocessed ECAP signals is 28.97% for $\bar{\varepsilon}_M$, showing that the noise resistance capability of the PECAP and PECAP-TSPD algorithms. The performance of the PECAP-TSPD algorithm is superior to that of the PECAP algorithm, as the values of $\bar{\varepsilon}_M$ and $\bar{\varepsilon}_A$ obtained with PECAP-TSPD are lower than those obtained with PECAP. The average of $\bar{\varepsilon}_M$ from Scenarios 1 to 7 can be ranked as: PECAP-TSPD (3.83%), PECAP (5.14%), and unprocessed (23.07%). The averages of $\bar{\varepsilon}_A$ from Scenarios 1 to 7 for PECAP-TSPD and PECAP are 3.01% and 4.64%, respectively.

Next, the impulse noise is added into the EACP matrix to evaluate the performance of PECAP-TSPD under four different densities. Figure 7 illustrates the results of the ECAP matrices before and after using the PECAP and PECAP-TSPD algorithms at the 40% density of the impulse noise.



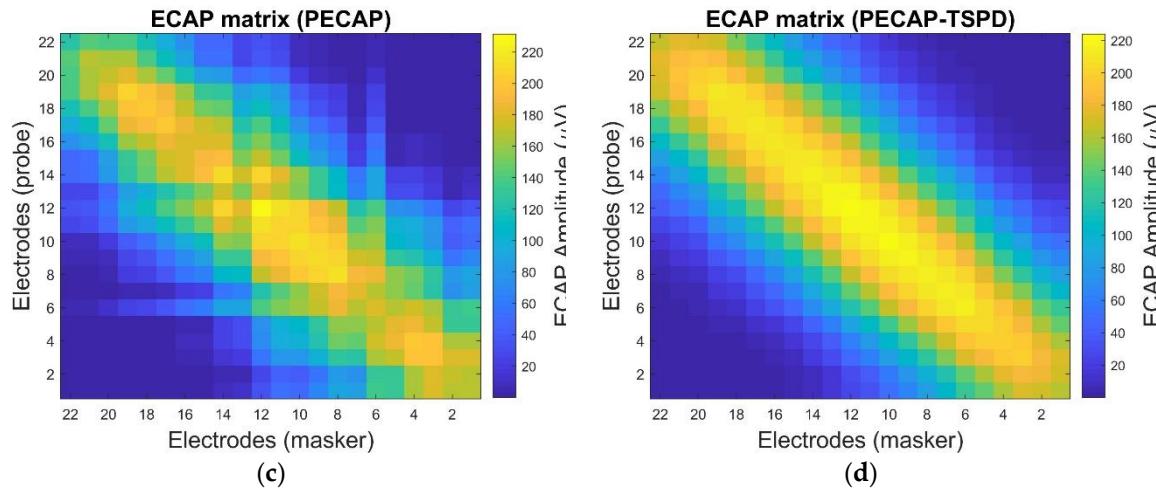


Figure 7. The ECAP matrix results of (a) non-density impulse noise, (b) impulse noise of 40% density, (c) processed ECAP matrix using the PECAP method under 40% density impulse noise, and (d) processed ECAP matrix using the PECAP-TSPD method under 40% density impulse noise. The settings of neural health and current spread are scenario 2 listed in Table 2.

The impulse noise with 40% density heavily contaminates the original ECAP matrix (Figure 7(a)), as shown in Figure 7(b), emphasizing the importance of signal processing. The processed ECAP matrices using the PECAP and PECAP-TSPD algorithms are depicted in Figures 7(c) and (d), where the impulse noise is most reduced. When comparing Figure 7(c) to Figure 7(d), the restored ECAP matrix in Figure 7(d) shows more resemblance to Figure 7(a) than to that of Figure 7(c), exhibiting the satisfactory performance of PECAP-TSPD under adverse noisy environments. The normalized RMSE results of the ECAP magnitude and neural activity pattern of the impulse noise case are depicted in Figure 8.

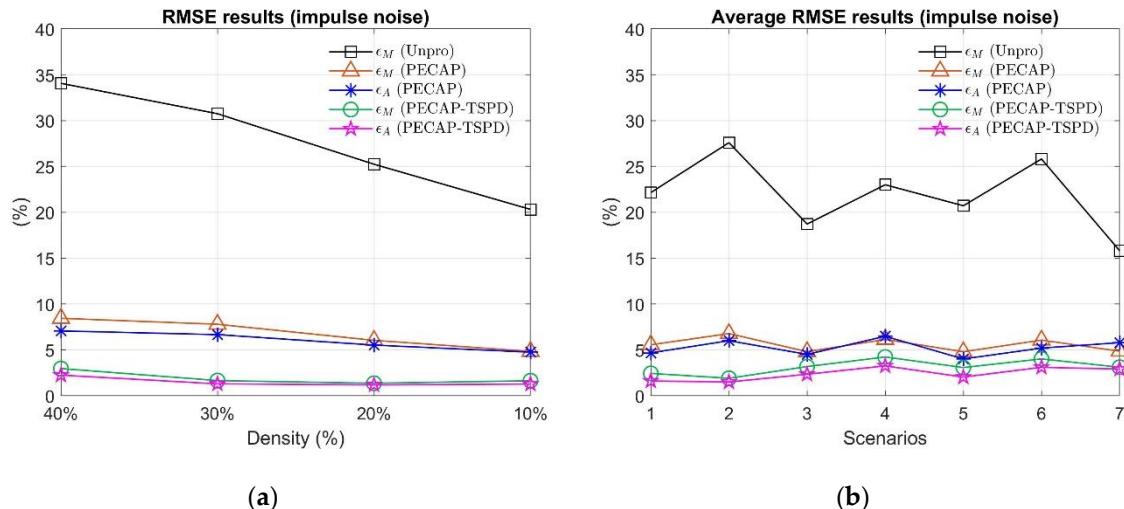


Figure 8. The normalized RMSE and average normalized RMSE results of ECAP magnitudes and neural activity patterns under four impulse noise densities (a) in scenario 2 and (b) in scenarios 1 to 7, respectively. The unprocessed ECAP matrices are compared with the PECAP and PECAP-TSPD algorithms. The settings for each scenario are listed in Table 2.

Figure 8(a) presents the ϵ_M and ϵ_A curves at four distinct densities in Scenario 2 for the unprocessed, PECAP, and PECAP-TSPD approaches, all of which increase as the impulse noise density increases. In the case of 40% impulse noise density, the ϵ_M values for unprocessed, PECAP, and PECAP-TSPD are 34.07%, 8.44%, and 2.96%, respectively, displaying the effectiveness of PECAP and PECAP-TSPD in highly noisy environments. The average normalized RMSE results are depicted

in Figure 6(b), in which $\bar{\varepsilon}_M$ and $\bar{\varepsilon}_A$ are similar to Eqs. (23) and (24), except that N_{SNR} is replaced by $N_{density} = 4$, denoting the total number of impulse noise densities in this work. The maximum value of $\bar{\varepsilon}_M$ for the PECAP and PECAP-TSPD algorithms is 6.77% and 4.22%, respectively. For the unprocessed ECAP matrices, the maximum value of $\bar{\varepsilon}_M$ is 27.59%, which suggests that PECAP and PECAP-TSPD are robust against noise. The PECAP-TSPD algorithm performs better than the PECAP algorithm because the values of $\bar{\varepsilon}_M$ and $\bar{\varepsilon}_A$ calculated by PECAP-TSPD are lower than those calculated by PECAP. The mean values of $\bar{\varepsilon}_M$ from Scenarios 1 through 7 can be arranged in ascending order as follows: PECAP-TSPD (3.13%), PECAP (5.57%), and unprocessed (21.97%). The mean values of $\bar{\varepsilon}_A$ from Scenarios 1 to 7 for PECAP-TSPD and PECAP are 2.39% and 5.23%, respectively.

5. Conclusions

The PECAP-TSPD algorithm, which integrates the improved median filter, the log-spectral Wiener filter, and PECAP, has been developed for noise reduction in ECAP data to accurately estimate neural health and current spread from severely noisy ECAP matrices. The normalized RMSE for ECAP magnitude (ε_M) and neural activity pattern (ε_A) demonstrates that both the PECAP and PECAP-TSPD algorithms effectively reduce ε_M and ε_A compared to unprocessed data under various SNRs, noise densities, and experimental scenarios. Notably, PECAP-TSPD outperforms PECAP in terms of both ε_M and ε_A . For ECAP matrices corrupted by random noise, the average ε_M across seven different scenarios and twelve SNR levels is as follows: PECAP-TSPD (3.83%), PECAP (5.14%), and unprocessed (23.07%). In the case of impulse noise, the ranking is: PECAP-TSPD (3.13%), PECAP (5.57%), and unprocessed (21.97%). Similarly, the average ε_A for PECAP-TSPD and PECAP under random noise conditions are 3.01% and 4.64%, respectively. Under impulse noise, the values are 2.39% for PECAP-TSPD and 5.23% for PECAP.

Author Contributions: Author Fan-Jie Kung designs the PECAP-TSPD algorithm, finishes the program, and writes this article.

Funding: This research was funded by the National Science and Technology Council (NSTC) of Taiwan, grant number 113-2222-E-027-010- and the APC was funded by 113-2222-E-027-010-.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: Author Fan-Jie Kung appreciates valuable comments from reviewers to improve this article.

Conflicts of Interest: Author Fan-Jie Kung has no conflicts to disclose.

Abbreviations

The following abbreviations are used in this manuscript:

ECAP	Electrically evoked compound action potential
PECAP	Panoramic ECAP
SNR	Signal-to-noise ratio
TSPD	Two-stage preprocessing denoising algorithm
LSA	Log-spectral amplitude
RMSE	Root mean square error
Unpro	Unprocessed data

References

1. Liebscher, T.; Hornung, J.; Hoppe, U. Electrically evoked compound action potentials in cochlear implant users with preoperative residual hearing. *Frontiers in Human Neuroscience* **2023**, *17*, 1-12. [[CrossRef](#)]

2. He, S.; Teagle, H. F. B.; Bunchman, C. A. The electrically evoked compound action potential: from laboratory to clinic. *Frontiers in Neuroscience* **2017**, *11*, 1-20. [[CrossRef](#)]
3. Hughes, M. L. Fundamentals of Clinical ECAP Measures in Cochlear Implants: Part 1: Use of the ECAP in Speech Processor Programming (2nd Ed.). Available online: [Fundamentals of Clinical ECAP Measures in Cochlear Implants: Part 1: Use of the ECAP in Speech Processor Programming \(2nd Ed.\)](#) (accessed on 13 February, 2025).
4. DeVries, L.; Schepeler, R.; Bierer, J. A. Assessing the electrode-neuron interface with the electrically evoked compound action potential, electrode position, and behavioral thresholds. *Journal of the Association for Research in Otolaryngology* **2016**, *22*, 237-252. [[CrossRef](#)]
5. Choi, C. T. M.; Wu, D. L. Electrically evoked compound action potential studies based on finite element and neuron models. *IEEE Transactions on magnetics* **2022**, *58*, 7501404. [[CrossRef](#)]
6. Garcia, C. The Panoramic ECAP Method: estimating patient-specific patterns of current spread and neural health in cochlear-implant users. PhD dissertation, University of Cambridge, 2 September, 2022. [[CrossRef](#)]
7. Dong, Y.; Briare, J. J.; Stronks, H. C.; Frijns, H.M. Speech perception performance in cochlear implant recipients correlates to the number and synchrony of excited auditory nerve fibers derived from electrically evoked compound action potentials. *Ear and Hearing* **2023**, *44*, 276-286. [[CrossRef](#)]
8. Takanen, M.; Strahl, S.; Schwarz, K. Insights into electrophysiological metrics of cochlear health in cochlear implant users using a computational model. *Journal of the Association for Research in Otolaryngology* **2024**, *25*, 63-78. [[CrossRef](#)]
9. Takanen, M.; Seeber, B. U. A phenomenological model reproducing temporal response characteristics of an electrically stimulated auditory nerve fiber. *Trends in Hearing* **2022**, *26*, 1-19. [[CrossRef](#)]
10. Garcia, C.; Deeks, J. M.; Goehring, T.; Borsetto, D.; Bance, M.; Carlyon, R. P. SpeedCAP: an efficient method for estimating neural activation patterns using electrically evoked compound action-potentials in cochlear implant users. *Ear and Hearing* **2023**, *44*, 627-640. [[CrossRef](#)]
11. Cosentino, S.; Gaudrain, E.; Deeks, J. M.; Carlyon, R. P. Multistage nonlinear optimization to recover neural activation patterns from evoked compound action potentials of cochlear implant users. *IEEE Trans. Biomedical Engineering* **2015**, *63*, 833- 840. [[CrossRef](#)]
12. Garcia, C.; Goehring, T.; Cosentino, S.; Turner, R. E.; Deeks, J. M.; Brochier, T.; Rughooputh, T.; Bance, M.; Carlyon, R. P. The panoramic ECAP method: estimating patient-specific patterns of current spread and neural health in cochlear implant users. *Journal of the Association for Research in Otolaryngology* **2021**, *22*, 567-589. [[CrossRef](#)]
13. Isnanto, R. R.; Windarto, Y. E.; Mangkuratmaja, M. V. Assessment on image quality changes as a result of implementing median filtering, Wiener filtering, histogram equalization, and hybrid methods on noisy images. In Proceedings of the International Conference on Information Technology, Computer, and Electrical Engineering, Semarang, Indonesia, 24-25 September, 2020. [[CrossRef](#)]
14. Gupta, G. Algorithm for image processing using improved median filter and comparison of mean, median and improved median filter. *International Journal of Soft Computing and Engineering* **2011**, *1*, 304-311. [[CrossRef](#)]
15. Sun, M. Comparison of processing results of median filter and mean filter on Gaussian noise. *Applied and Computational Engineering* **2023**, *5*, 779-785. [[CrossRef](#)]
16. Hou, Y.; Li, Q.; Zhang, C.; Lu, G.; Ye, Z.; Chen, Y.; Wang, L.; Cao, D. The state-of-the-art review on applications of intrusive sensing, image processing techniques, and machine learning methods in pavement monitoring and analysis. *Engineering* **2021**, *7*, 845-856. [[CrossRef](#)]
17. Zhu, Y.; Huang, C. An improved median filtering algorithm for image noise reduction. *Physics Procedia* **2021**, *25*, 609-616. [[CrossRef](#)]
18. Jiang, D. A study on adaptive filtering for noise and echo cancellation. Master's thesis, University of Windsor, 2005. [[CrossRef](#)]
19. Yazdanpanah, H.; Diniz, P. S. R. Recursive least-squares algorithms for sparse system modeling. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, New Orleans, LA, USA, 5-9 March, 2017. [[CrossRef](#)]

20. Creighton, J.; Doraiswami, R. Real time implementation of an adaptive filter for speech enhancement. In Proceedings of the Canadian Conference on Electrical and Computer Engineering, Niagara Falls, ON, Canada, 2-5 May, 2004. [[CrossRef](#)]
21. Wang, P.; Kam, P.-Y. An automatic step-size adjustment algorithm for LMS adaptive filters, and an application to channel estimation. *Physical Communication* **2012**, *5*, 280-286. [[CrossRef](#)]
22. Loizou, P. C. *Speech Enhancement: Theory and Practice*, 1st ed.; CRC Press: New York, USA, 2007; pp. 143–208. [[CrossRef](#)]
23. Benesty, J.; Chen, J.; Huang, Y. *Microphone Array Signal Processing*, Springer: Verlag Berlin Heidelberg, Germany, 2008; pp. 8–15. [[CrossRef](#)]
24. Ephraim, Y.; Malah, D. Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. *IEEE Trans. Acoustics, Speech, and Signal Processing* **1985**, ASSP-33, 443-445. [[CrossRef](#)]
25. Borgström, B. J.; Alwan, A. Log-spectral amplitude estimation with Generalized Gamma distributions for speech enhancement. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Czech Republic, 22-27 May, 2011. [[CrossRef](#)]
26. Hirszhorn, A.; Dov, D.; Talmor, R.; Cohen, I. Transient interference suppression in speech signals based on the OM-LSA algorithm. In Proceedings of the International Workshop on Acoustic Echo and Noise Control, Aachen, Germany, 4-6 September 4-6, 2012. [[CrossRef](#)]
27. Hsu, Y.; Bai, M. R. Learning-based robust speaker counting and separation with the aid of spatial coherence. *Eurasip Journal on Audio, Speech, and Music Processing* **2023**, *2023*, 1-21. [[CrossRef](#)]
28. Hsu, Y.; Lee, Y.; Bai, M .R. Array configuration-agnostic personalized speech enhancement using long-short-term spatial coherence. *Journal of the Acoustical Society of America* **2023**, *154*, 2499-2511. [[CrossRef](#)]
29. Richard, G.; Smaragdis, P.; Gannot, S.; Naylor, P. A.; Makino, S.; Kellermann, W.; Sugiyama, A. Audio signal processing in the 21st century: The important outcomes of the past 25 years. *IEEE Signal Processing Magazine* **2023**, *40*, 12-26. [[CrossRef](#)]
30. Gannot, S.; Tan, Z.-H.; Haardt, M.; Chen, N. F.; Wai, H.-T.; Tashev, I.; Kellermann, W.; Dauwels, J. Data science education: the signal processing perspective. *IEEE Signal Processing Magazine* **2023**, *40*, 89-93. [[CrossRef](#)]
31. Tan, K.; Wang, D. A convolutional recurrent neural network for real-time speech enhancement. In Proceedings of the Interspeech, Hyderabad, India, 2-6 September, 2018. [[CrossRef](#)]
32. Bianco, M. J.; Gerstoft, P.; Traer, J.; Ozanich, E.; Roch, M. A.; Gannot, S.; Deledalle, C.-A. Machine learning in acoustics: theory and applications. *Journal of the Acoustical Society of America* **2019**, *146*, 3590-3628. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.