This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TETCI.2025.3577504, IEEE Transactions on Emerging Topics in Computational Intelligence

Adaptive Thresholding in EEG Artifact Removal through Multimodal Fusion: A Multimodal Artifact Subspace Reconstruction Approach

Wenlong You, Rui Yang, *Senior Member, IEEE*, Chengxuan Qin, Mengjie Huang, Zidong Wang, *Fellow, IEEE*

***Abstract*—The removal of artifacts is essential for improving the quality and reliability of electroencephalogram (EEG) data in academic research. Traditional methods, such as blind source separation and signal space projection, often involve subjective and time-consuming manual parameter selection, which is ineffective for artifacts closely correlated with EEG signals. Furthermore, existing artifact removal methods are difficult to generalize across different datasets and experimental conditions. Although artifact subspace reconstruction shows promise, it remains computationally complex and sensitive to parameter selection, limiting its real-time applicability and ability to handle complex artifacts. This study proposes the Multimodal Artifact Subspace Reconstruction (MASR) method, which reduces manual intervention and improves automatic detection and removal of complex artifacts. MASR proposes a new use of multimodal feature extraction techniques, innovatively providing an informative reference for processing EEG signals to reduce artifacts across channels. MASR enhances artifact removal by introducing a novel channel significance metric for quantifying artifact contamination and employing a dynamic adaptive threshold to reduce parameter dependency. MASR integrates multimodal features through principal component analysis (PCA) and ensures cross-modal consistency with Pearson correlation coefficient (PCC) for EEG artifact removal, solving the challenge of artifact characteristics. The MASR method offers a robust, data-driven solution that improves the quality and reliability of EEG data across various applications.**

***Index Terms*—Artifact removal, artifact subspace reconstruction, brain-computer interface, electroencephalogram, multimodality fusion, transfer spectral entropy.**

I. Introduction

BRAIN-COMPUTER interface (BCI) systems enable users to communicate directly with a computer using brain signals [42], and can leverage various modalities as control signals, such as functional magnetic resonance imaging (fMRI), positron emission tomography (PET), and electroencephalogram (EEG) [26], [27], [38]. Among these modalities, EEG is the most commonly used in BCI due to its exceptional temporal resolution, portability, and reliability [39]. EEG measures the joint electrical activity of a population of neurons, typically with an amplitude of a few microvolts, which can be extracted and decoded for specific tasks in the BCI system [10]. The process of collecting EEG data involves recording the electrical activity of brains through electrodes placed on the scalp. However, the recording process often introduces unwanted noise from various artifacts, which can distort the original brain activity and complicate the analysis of EEG signals [25].

The successful implementation of EEG in both clinical and industrial environments relies on the effective removal of artifacts. Artifacts are considered interference or noise components in brain signals and arise from a variety of sources, including but not limited to environmental factors [26], instrument limitations [16], and physiological activities (such as eye blinks and muscle contractions) [21]. The development of artifact removal methods is essential to address noise interference and ensure precise interpretation of EEG data, highlighting the need for innovation in artifact detection and mitigation methods [26]. Recent literature reviews indicate a growing focus on the development of artifact removal methods aimed at eliminating non-brain components from recorded EEG signals [7], [14].

In EEG analysis, a prominent method for artifact removal refers to blind source separation (BSS) and particularly independent component analysis (ICA). ICA can reconstruct EEG signals by eliminating artifact sources when these sources are statistically independent and non-Gaussian [12], [14], [24], [47]. XDAWN is a spatial filtering method designed to improve the signal-to-signal and noise ratio of the evoked response [31]. However, such dependence on the *characteristics of artifacts* often necessitates combination with other decomposition methods [14]. SSP effectively mitigates noise by projecting data into a subspace orthogonal to the noise direction, provided that the noise direction in the signal space is well-defined and stable [26], [37]. Nevertheless, the limitations of SSP stem from its reliance on specific *channel configurations*. Recent advancements in EEG analysis have introduced artifact subspace reconstruction (ASR) as a contemporary method designed to remove complex artifacts from EEG

This work has been approved by University Ethics Committee of Xi'an Jiaotong-Liverpool University with proposal number EXT20-01-07 on March 31 2020, and is partially supported by National Natural Science Foundation of China (72401233), Jiangsu Provincial Qinglan Project, Natural Science Foundation of Jiangsu Higher Education Institutions of China (23KJB520038), and Research Enhancement Fund of XJTLU (REF-23-01-008).

W. You and C. Qin are with the School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, 215123, China, and also with the School of Electrical Engineering, Electronics and Computer Science, University of Liverpool, Liverpool, L69 3BX, United Kingdom (e-mail: Wenlong.You23@student.xjtlu.edu.cn; C.Qin8@liverpool.ac.uk);

M. Huang is with Design School, Xi'an Jiaotong-Liverpool University, Suzhou, 215123, China (e-mail: Mengjie.Huang@xjtlu.edu.cn);

R. Yang is with the School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, 215123, China (e-mail: R.Yang@xjtlu.edu.cn).

Z. Wang is with the Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, United Kingdom (e-mail: Zidong.Wang@brunel.ac.uk).

Corresponding authors: R. Yang.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TETCI.2025.3577504, IEEE Transactions on Emerging Topics in Computational Intelligence

signals autonomously [3], [8]. ASR excels at autonomously removing large-amplitude artifacts, particularly in contexts with substantial EEG variability [40]. Dynamic modulation of ASR parameters could improve artifact removal and signal preservation [8], [13]. However, a limitation of ASR is its *parameter dependency*, as it identifies artifact-contaminated segments based on thresholds derived from baseline EEG data. Variability in baseline EEG and unpredictable noise make stable reference extraction challenging, necessitating additional EEG-derived information to ensure reliable artifact removal [4]. Common spatial pattern (CSP) and convolutional neural network (CNN)-based methods have shown promise in artifact removal [26]. However, their performance often depends on the availability of large, labeled datasets for training, which may not always be feasible.

The limitations of artifact removal methods arise partly due to the complexity and high dimensionality of signals, which complicates the development of approaches capable of adapting to a wide range of populations [2]. Nevertheless, the extensive adoption of wearable devices and progress in data acquisition and analysis has not only established a foundation but also highlighted the demand for innovative multimodal approaches capable of efficiently handling and integrating multimodal information from multiple sources [23], [32]. These multimodal approaches enhance EEG analysis by incorporating additional information for EEG analysis, thus providing representative features that aid in improving artifact removal [34], [44]. In the field of BCI, multimodal data refers to inputs from various devices or feature representation (utilizing distinct feature extraction techniques to derive different types of features). In such a context, every input or feature representation can be regarded as one modality [46]. Seamlessly leveraging multimodal approaches for artifact analysis involves two primary challenges: ensuring cross-modal consistency and effectively integrating multimodal information.

Cross-modal consistency ensures that representations across different modalities within a shared feature space convey the same underlying meaning [43]. Methods based on fusion are generally integrated with feature extraction methods and aim to minimize redundant information. Typical methods for feature extraction include principal component analysis (PCA) [35] and max-relevance and min-redundancy (mRMR) [18]. Modal information extracted from each feature representation of EEG signals can also be leveraged by multimodal fusion methods [5], [17], [28], [29]. EEG signals can extract various types of modal information, such as time-frequency, frequency-domain, and time-domain information. Multimodal approaches offer solutions by combining various data sources and analytical techniques to effectively distinguish genuine brain signals from artifacts [41], [45], [46].

This study focuses on the improvement of existing methods to accommodate automated detection and removal of complex artifacts from EEG data by leveraging the integration of multimodal information. Specifically, a multimodal artifact subspace reconstruction (MASR) method is proposed. To ensure cross-modal consistency, the study employs a standardization process that aligns information from different modalities, ensuring that the data from different modalities represent equivalent meanings within a shared feature space. This process addresses the inherent gaps between modalities, facilitating the integration of multimodal information for artifact detection. The study includes a comparative evaluation of MASR with ICA, SSP, and traditional ASR methods, demonstrating that MASR outperforms these methods in artifact suppression and overall performance. The main contributions of this paper can be summarized as follows:

1) MASR introduces a metric called channel significance to quantify artifact pollution in EEG data by evaluating the impact of artifacts on individual channels, facilitating a stratification mechanism that tackles shortcomings *channel configurations*;
2) MASR employs a dynamic adaptive threshold to ascertain the cutoff parameter in ASR methods, predicated on a coefficient of variation, thereby effectively addressing *parameter dependency* and improving the accuracy of artifact removal while minimizing the loss of critical EEG information;
3) MASR enhances EEG artifact removal by aligning multimodal features into a shared space and fusing them using PCA, thereby overcoming challenges in cross-modal consistency and effective multimodal integration imposed by *characteristics of artifacts*.

The remaining sections of this paper are as follows: Section II describes the proposed MASR method; Section III presents the experimental setting and offers an in-depth analysis of the results; finally, the conclusions and discussions on relevant future work are presented in Section IV.

II. Multimodal Artifact Subspace Reconstruction

The proposed MASR method consists of three main modules: the multimodal feature extraction module, the significance-based channel grouping module, and the module of adaptive thresholding informed by coefficient of variation, as shown in Figure 1. Following the segmentation of EEG data, the multimodal feature extraction module initiates its operation by simultaneously extracting and computing feature information. This module integrates multimodal information from transfer spectral entropy (TSE) [9], power spectral density (PSD) [11], and wavelet packet decomposition (WPD) [19] to capture diverse artifact characteristics, providing focused data segments for subsequent analysis.

Upon completion of the segmentation, the channel grouping module classifies the EEG channels based on their channel significance using statistical metrics, thereby forming channel groups for targeted artifact removal. These groups are then processed in the subsequent module, which applies the coefficient of variation to compute adaptive thresholds. This ensures that these thresholds are dynamically adjusted based on the statistical variability of the channel significance scores within each channel group, optimizing artifact removal while preserving the integrity of the EEG signal. In this framework, artifact removal techniques tailored to the unique characteristics of each group can complement each other, maximizing artifact suppression while minimizing the loss of EEG information.

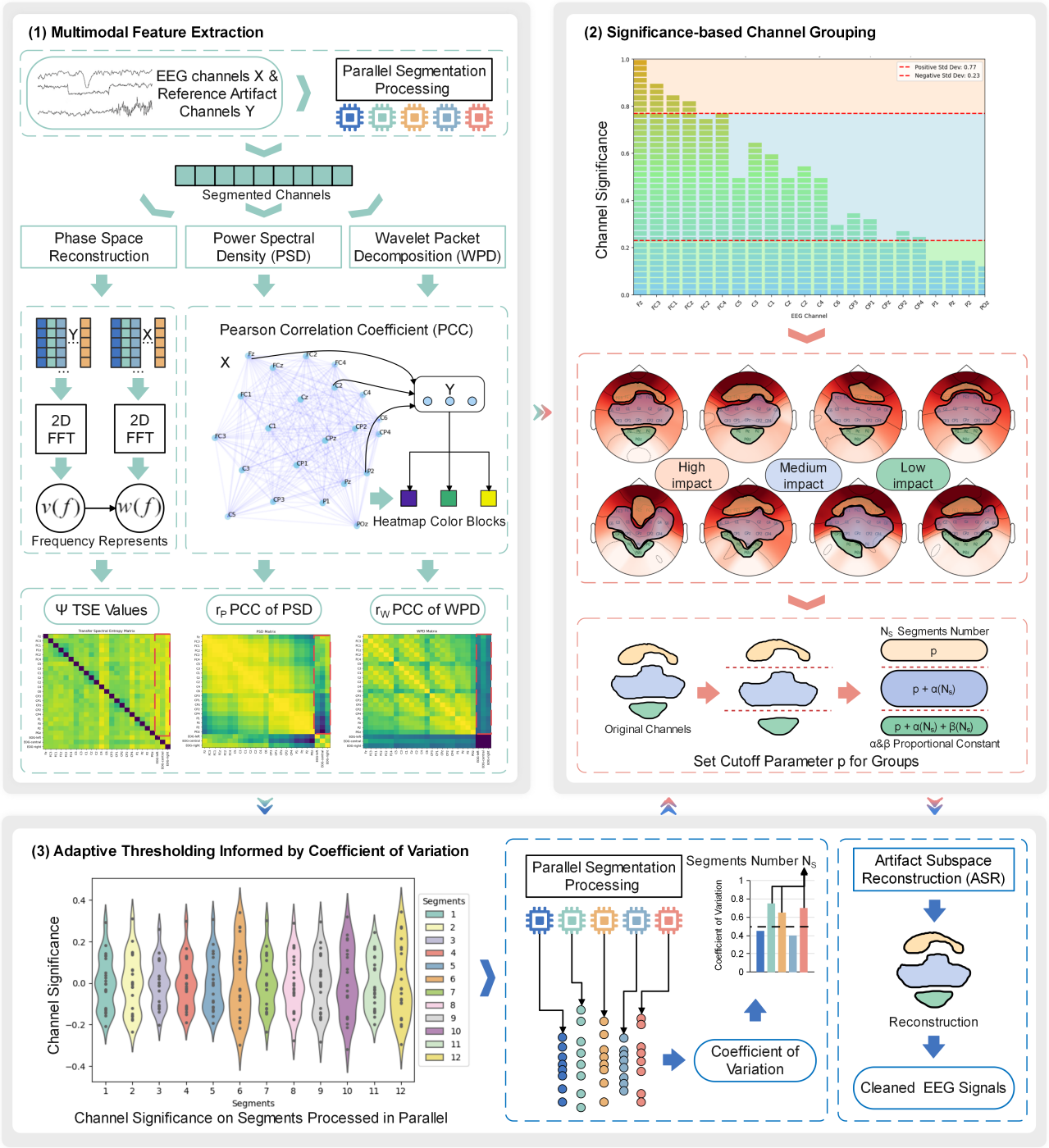


Fig. 1. The flow diagram of MASR method

A. Multimodal Feature Extraction

In this phase, the impact of artifacts on EEG data is quantified using a multimodal fusion strategy integrating three sources of modal information: TSE, WPD, and PSD, as illustrated in part (1) of Figure 1. To ensure *cross-modal consistency*, MASR standardizes the feature representations from modal information, aligning them within a shared feature space. The subsequent sections detail the computational methodologies employed for the three modal information and the execution of *cross-modal consistency* mechanisms. The data is processed in parallel operations in segments, and then matrices Ψ, r_{P}, and r_{W} are horizontally stacked to form a combined feature matrix F, representing the interactions between EEG and artifact reference channels across multiple

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TETCI.2025.3577504, IEEE Transactions on Emerging Topics in Computational Intelligence

modalities as follows (Table I shows the list of symbols with corresponding descriptions):

$$\mathbf{F} = \begin{bmatrix} \Psi & r_{\mathrm{P}} & r_{\mathrm{W}} \end{bmatrix} \tag{1}$$

TABLE I
LIST OF SYMBOLS

Symbol	Description
α	proportional constant of mapping function
β	proportional constant of mapping function
δ	coefficient of variation
λ	channel significance
μ	mean value of the channel significance
σ	standard deviation value of the channel significance
φ	dimension of the delay vector
ϕ	dimension of time points within the delay vector
$\Psi_{Y\to X}$	TSE from Y to X
$1(\bullet)$	indicator function
$d_X(i)$	WPD of X at time index i
$\overline{d_X}$	mean WPD value of X
$d_Y(i)$	WPD of Y at time index i
$\overline{d_Y}$	mean WPD value of Y
$\mathbf{F}$	multimodal feature matrix
f	frequency index derived from sampling frequency
G_i	channel group of category i
$H(\bullet)$	entropy
N	number of EEG channels
N_s	number of segments exceeded the threshold
n	number of segments created for each group
r_P	Pearson correlation coefficient of PSD
r_W	Pearson correlation coefficient of WPD
p	cutoff parameter of ASR
$\mathbb{R}$	represents the real number of space
$S_X(i)$	PSD of X at frequency index i
$\overline{S_X}$	mean PSD value of X
$S_Y(i)$	PSD of Y at frequency index i
$\overline{S_Y}$	mean PSD value of Y
s	segment of signal channels
T	threshold of coefficient of variation
$v(f)$	vector of elements from Y
$v^{\varphi}(f)$	φ-dimensional delay vector for vector $v(f)$
$\mathbf{w}$	first principal component
$w(f)$	vector of elements from X
$w^{\phi}(f)$	ϕ-dimensional delay vector for vector $w(f)$
X	EEG channels
Y	reference artifact channels

- Transfer Spectral Entropy (TSE) is an extension of transfer entropy designed to investigate local frequency band coupling in the frequency domain [9]. Firstly, the raw data undergoes phase space reconstruction [6]. Subsequently, a 2-dimensional fast Fourier transform (2D FFT) is performed to obtain the frequency represents $w(f)$ and $v(f)$ in the frequency domain. Finally, transfer entropy values are computed for each pair of $w(f)$ and $v(f)$ with the same frequency index to derive TSE value Ψ from reference artifact channels Y to EEG channels X as follows:

$$\begin{aligned} \Psi_{Y\to X}(f) = {} & H(v(f) \mid v^{\varphi}(f)) \\ & - H(v(f) \mid w^{\phi}(f), v^{\varphi}(f)) \end{aligned} \tag{2}$$

where $v^{\varphi}(f)$ and $w^{\phi}(f)$ represent the φ- and ϕ-dimensional delay frequency representations respectively, $H(v(f) \mid v^{\varphi}(f))$ denotes the conditional entropy of the processed $v(f)$ given its past values, and $H(v(f) \mid w^{\phi}(f), v^{\varphi}(f))$ denotes the conditional entropy of the processed $v(f)$ given both $w^{\phi}(f)$ and its past values. φ is defined by the number of past time points considered, and the index of the time points within the delay vector is represented by ϕ. TSE serves as a quantitative measure to evaluate the influence of artifacts on EEG signals across different channels.
- Power Spectral Density (PSD) enables the analysis of the spectrum of time domain signals and the capture of energy distribution at different frequencies [11], [30]. Pearson correlation coefficient (PCC) helps to assess the degree of relationship between the artifact reference channels and the EEG channels, transforming the feature information in the frequency domain of the signals into correlations between the signals, thereby achieving *cross-modal consistency* in the task of artifact removal. Then PCC is computed across the PSD of channels X and Y, denoted as r_P below:

$$r_P = \frac{\sum_{i=1}^{n}(S_X(i) - \overline{S_X})(S_Y(i) - \overline{S_Y})}{\sqrt{\sum_{i=1}^{n}(S_X(i) - \overline{S_X})^2}\sqrt{\sum_{i=1}^{n}(S_Y(i) - \overline{S_Y})^2}} \tag{3}$$

where $S_X(i)$ and $S_Y(i)$ represent the estimated PSD values of EEG channel X and reference artifact channel Y at frequency index i, $\overline{S_X}$ and $\overline{S_Y}$ represent the mean values of these estimated PSD, and n denote the total number of frequency indices considered.
- Wavelet Packet Decomposition (WPD) can be regarded as a continuous-time wavelet decomposition sampled at different frequencies at every level or scale [19], [33]. PCC measures the linear correlation between the two sets of data, with values ranging from -1 (perfect negative correlation) to 1 (perfect positive correlation). The PCC-based process in the MASR represents an innovation, enhancing the artifact removal accuracy through a more refined *cross-modal consistency* evaluation, allowing for optimized correlation analysis between EEG and reference artifact channels. Similar to the case with PSD, WPD also does not directly capture the relationship between artifact reference channels and EEG channels. Therefore, calculating the PCC is also necessary for WPD, with the PCC for the WPD of signals X and Y denoted as r_W below:

$$r_W = \frac{\sum_{i=1}^{n}(d_X(i) - \overline{d_X})(d_Y(i) - \overline{d_Y})}{\sqrt{\sum_{i=1}^{n}(d_X(i) - \overline{d_X})^2}\sqrt{\sum_{i=1}^{n}(d_Y(i) - \overline{d_Y})^2}} \tag{4}$$

where $d_X(i)$ and $d_Y(i)$ represent the WPD values of EEG channel X and reference artifact channel Y at time index i, $\overline{d_X}$ and $\overline{d_Y}$ represent the mean values of these WPD, and n denotes the total number of time indices considered.

Remark 1. *The number of data segments can generally be chosen based on the available number of CPU cores, with a multiple of the core count being recommended for optimal performance. In the experimental setup of this work, the data was divided into twelve segments for parallel processing, ensuring the best performance for data processing. In different*

environments, the number of data segments can be adjusted to a higher or lower value as needed.

B. Significance-based Channel Grouping

The fusion of multimodal features and the proposal of channel significance aim to address the lack of attention given to the targeted, personalized treatment of channels in existing approaches. The significance-based channel grouping module corresponds to part (2) of Figure 1.

To extract the dominant variability across channels, MASR applies PCA to the multimodal feature matrix $\mathbf{F} \in \mathbb{R}^{N\times M}$, where $\mathbb{R}$ represents the real number space. The first principal component $\mathbf{w}$ is obtained as the eigenvector corresponding to the largest eigenvalue. Then channel significance λ_i is then defined as:

$$\lambda_i = \frac{\left|\mathbf{F}_{i,:}^{\top}\mathbf{w}\right|}{\max\limits_{1\le j\le N}\left|\mathbf{F}_{j,:}^{\top}\mathbf{w}\right|}, \quad i = 1, 2, \ldots, N. \tag{5}$$

where λ_i represents the relative contribution of channel i to the primary mode of variation in $\mathbf{F}$, effectively quantifying its significance for artifact removal. This definition ensures that λ_i is normalized across all channels, so the values lie within the range $[0, 1]$. Channels are then grouped based on statistical thresholds derived from the normalized λ_i.

Based on the channel significance $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_N]$, where N represents the total number of EEG channels, the channels are grouped into three categories based on their significance to facilitate targeted artifact removal. The mean μ and standard deviation σ of the significance scores are computed as $\mu = \frac{1}{N}\sum_{i=1}^{N}\lambda_i$, and $\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\lambda_i - \mu)^2}$, where N is the total number of EEG channels and λ_i represents the channel significance score for channel i. The grouping process begins with the calculation of statistical thresholds. Two thresholds can be established: the low threshold $T_l = \mu - \sigma$, and the high threshold $T_h = \mu + \sigma$. Each EEG channel is then allocated to one of three groups (denoted as G_{low}, G_{medium}, and G_{high}) based on its significance score λ_i relative to the thresholds:

$$\begin{aligned} G_i =& \mathbf{1}(\lambda_i > T_h)\cdot G_{high} + \mathbf{1}(\lambda_i < T_l)\cdot G_{low} \\ &+ \mathbf{1}(T_l \le \lambda_i \le T_h)\cdot G_{medium} \end{aligned} \tag{6}$$

where $\mathbf{1}(\lambda_i > T_h)$, $\mathbf{1}(\lambda_i < T_l)$ and $\mathbf{1}(T_l \le \lambda_i \le T_h)$ are indicator functions that assign a value of 1 when the respective conditions are satisfied, and 0 otherwise. Consequently, MASR assigns each channel to one of three groups based on the corresponding significance score λ_i and the thresholds T_l and T_h. The resulting G_i stores the grouping information and is passed to the next module.

C. Adaptive Thresholding Informed by Coefficient of Variation

MASR employs threshold-based adaptive grouping parameter assignment to prioritize artifact removal accuracy and minimize loss of EEG information during ASR processing. Part (3) of Figure 1 illustrates the flow of adaptive thresholding informed by the coefficient of variation. With the EEG data partitioned into n segments, the coefficient of variation δ was computed to assess the relative dispersion in the channel significance λ for each segment, thereby reflecting the degree of variability within each segment. δ_s represents the coefficient of variation of the significance scores within each segment s ($s \in \{1, 2, \ldots, n\}$), defined as the ratio of the standard deviation σ_s to the mean μ_s: $\delta_s = \sigma_s/\mu_s$. A high coefficient of variation indicates significant variability in the features, suggesting the need for parameter adjustment in the segmentation process, which is performed in parallel. To further quantify the variability, N_s represents the total count of segments for which the coefficient of variation δ_s exceeds the threshold T were tracked as:

$$N_s = \sum_{s=1}^{n}\mathbf{1}(\delta_s > T) \tag{7}$$

where $\mathbf{1}(\delta_s > T)$ is the indicator function. Based on the group assignment G_{low}, G_{medium}, and G_{high} from the preceding module, the cutoff parameter p was adaptively adjusted. The grouping G_i reflects the channel significance, and the parameter p_i for each channel was set as follows:

$$\begin{aligned} p_i = & p\cdot\mathbf{1}(G_i = G_{high}) \\ & + (p + \alpha(N_s))\cdot\mathbf{1}(G_i = G_{medium}) \\ & + (p + \alpha(N_s) + \beta(N_s))\cdot\mathbf{1}(G_i = G_{low}) \end{aligned} \tag{8}$$

where G_i denotes the group assignment for a given channel, α and β are proportional constants, and N_s quantifies the variability of segments exceeding the threshold T. The selection of α and β can be defined to adaptively adjust to the number of data segments, ensuring minimal differences in the cutoff parameter across groups and maintaining values within an appropriate range. This setup ensures that parameters are adjusted to provide differentiated artifact removal strategies tailored to the significance of each channel group. MASR's adaptive parameter selection eliminates the need for compromise in settings across all channels, enabling users to focus on artifact-affected channels for more precise removal.

By addressing the parameter problem of ASR for each channel group, MASR applies an artifact removal process to reconstruct the cleaned EEG data. The EEG channels are grouped based on channel significance, with adaptive thresholding set according to the coefficient of variation from segmented statistics, ensuring that the final output retains the significant features of the original EEG data while effectively minimizing artifact interference. Algorithm 1 outlines the algorithm flow of the MASR method.

III. Experimental Design and Results Analysis

A. Experimental Design

In the experiment, multi-modal feature extraction methods (TSE, PSD, and WPD) were applied to extract features. Correlation matrices for TSE, PSD, and WPD were generated to evaluate interactions between different channels. The MASR method processed the multimodal information to remove artifacts, operating at three channel groups: low-impact, medium-impact, and high-impact. To assess the effectiveness of the MASR method, comparisons were made with three other

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TETCI.2025.3577504, IEEE Transactions on Emerging Topics in Computational Intelligence

Algorithm 1: Multimodal Artifact Subspace Reconstruction (MASR)

Input : EEG channels $\mathbf{X}$, reference artifact channels $\mathbf{Y}$
Output: Cleaned EEG signals

for *each segment* s *of* X *and* Y **do**
  Compute TSE $\Psi_{Y\to X}$ from the conditional entropy of frequency segments using (2);
  Estimate Welch's method PSD and compute r_P as the PCC between X and Y using (3);
  Compute WPD coefficients and compute r_W as the PCC between X and Y using (4);
  Stack matrix of Ψ, r_P, and r_W to matrix $\mathbf{F}$ by (1).
Obtain channel significance by applying PCA on $\mathbf{F}$ using (5);
Average the channel significance from each segment n to obtain the overall channel significance λ;
Compute mean μ and standard deviation σ of the channel significance λ_i;
Establish thresholds: $T_l = \mu - \sigma$, $T_h = \mu + \sigma$;
Establish groups G_{low}, G_{medium}, and G_{high};
Assign channels to the groups using (6);
Initialize N_s and p;
for *each segment* s *in* $\{1, 2, \ldots, n\}$ **do**
  Compute μ_s and σ_s of channel significance δ_s;
  Compute δ_s and increment N_s by using (7).
Define cutoff parameter p_i for each group G_i using (8);
Apply artifact removal process to channel groups with the cutoff parameter p_i and then reconstruct;
return Cleaned EEG signals;

artifact removal methods: ICA, SSP, and ASR. To ensure fairness in the comparison, all methods were set to use the same eye movement channels as references. The effectiveness of the ASR was demonstrated by comparing the processed data with the original data through different waveforms. Finally, the processed data were analyzed using four experimental metrics to evaluate the effectiveness of the MASR method. The ablation experiment systematically assessed the impact of different modality combinations by evaluating the performance of single, dual, and triple modality configurations on signal accuracy and robustness across subjects.

1) Experimental Dataset: The performance of multimodal artifact subspace reconstruction is assessed using the BCI Competition IV 2a (BCIIV-2a) dataset, an openly accessible and widely recognized motor imagery-based EEG (MI-EEG) dataset [36]. Serving as a benchmark dataset for MI-EEG classification tasks, the BCIIV-2a dataset contains signals captured under uncontrolled conditions, making it particularly challenging due to the presence of various artifacts. During each experiment, nine participants were instructed to perform one of four motor imagery tasks: 1) imagining movements of the left hand; 2) right hand; 3) feet; and 4) tongue. Each participant underwent two recording sessions on separate days, with each session comprising 288 trials, resulting in a total of 576 trials per participant. In total, the dataset includes 5184 trials, where each trial represents a single sample in MI-EEG tasks. The recordings were conducted using 22 Ag/AgCl EEG electrodes along with three monopolar electrooculogram (EOG) electrodes. To estimate the impact of EOG interference, a 5-minute baseline recording was taken at the start of each trial. Based on both the literature review and experimental tests, we observed that applying artifact removal methods to high-quality EEG data can inadvertently lead to the loss of valuable neural information [26]. Consequently, the third subject, which demonstrated exceptionally high quality and negligible artifacts, was excluded from the artifact removal task, while the remaining eight subjects with pronounced artifacts were selected for the experiments.

2) Evaluation Metrics: In this study, four key metrics, including root mean square error (RMSE), normalized mean square error (NMSE), signal-to-artifact ratio (SAR), and mutual information (MI) are used to evaluate the performance of the proposed algorithm. These metrics assess the accuracy and signal quality following artifact removal, which are also used in other EEG artifact removal studies for performance evaluation, as discussed in previous research [25]. A summary of these metrics is provided in Table II, where X_t represents the signal before artifact removal, and Y_t represents the signal after artifact removal. Lower RMSE and NMSE values indicate reduced reconstruction errors, whereas higher SAR and MI values correspond to improved signal fidelity and more effective artifact suppression.

3) Parameter and Channel Settings: In the MASR method, EEG channels are classified into low, medium, and high-impact categories based on transformed features, using statistical thresholds determined by the mean and standard deviation. The initial cutoff parameter p is uniformly set to 20 for all channel groups. Given that the number of segments is fixed at 12, the values of α and β are both set to 0.5, ensuring a balanced adjustment during the adaptive thresholding process.

To justify these settings, prior studies on ASR suggest that a cutoff between 20 and 30 is effective for EEG signals, making 20 a reasonable choice [3]. The segmentation number of 12 was selected to match the 12 cores of our processing machine, thereby enhancing parallel processing efficiency. With α and β set to 0.5, the effective parameter range is maintained within an appropriate window for artifact removal. These settings provide a consistent basis for the experiments while allowing adjustments based on computational resources.

Based on the channel significance evaluation in Fig. 2,

TABLE II
METRICS FOR EVALUATING METHOD PERFORMANCE

Metrics	Formula
RMSE	$\sqrt{\frac{\sum_{t=1}^{n}(X_t - Y_t)^2}{n}}$
NMSE	$\frac{\sum_{t=1}^{n}(X_t - Y_t)^2}{\sum_{t=1}^{n} Y_t^2}$
SAR	$10\log_{10}\left(\frac{\sum_{t=1}^{n} X_t^2}{\sum_{t=1}^{n}(X_t - Y_t)^2}\right)$
MI	$\iint f(X_t, Y_t)\log\left(\frac{f(X_t,Y_t)}{f(X_t)f(Y_t)}\right) dX_t\, dY_t$

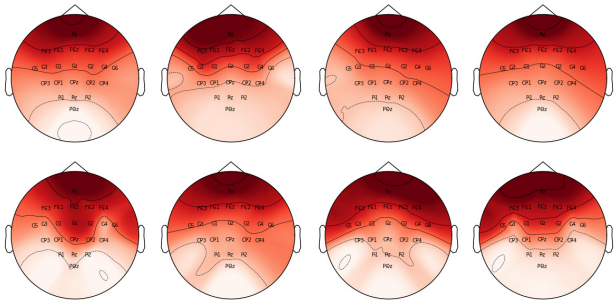


Fig. 2. Channel significance across subject

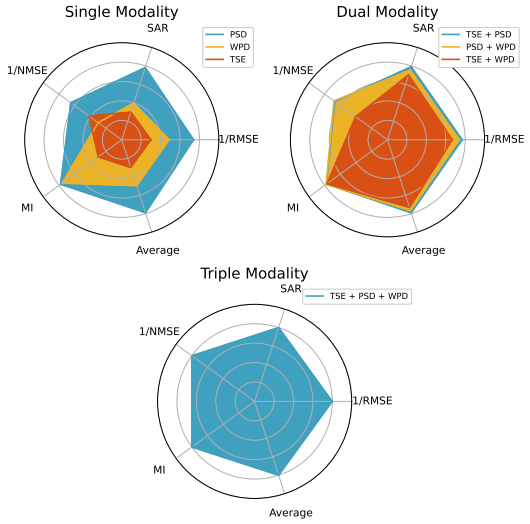


Fig. 3. Radar diagram of ablation modalities

EEG channels are adaptively assigned to high, medium, and low-impact groups according to their sensitivity to artifacts, as illustrated in Table III. The combined analysis of our figures, tables, and the intrinsic characteristics of the BCIIV-2a dataset demonstrates that the distinct features between artifacts and EEG signals are effectively captured and quantified by the proposed metrics. Furthermore, since the PCC inherently provides a standardized measure, only channel significance normalization was applied without any additional adjustments.

B. Ablation Experiment

Table IV summarizes the results of the modality ablation experiments, with data averaged across eight subjects. In the single-modality tests, the PSD modality achieved the lowest RMSE of 2.95×10^{-6} and a commendable SAR of 15.63, indicating that PSD alone effectively captures essential signal features. In the dual-modality assessments, the integration of TSE and PSD attained superior results, while the NMSE decreased to 0.0661. This suggests a synergistic effect, enhancing overall performance through the integration of complementary features. This integration also showed a notably high SAR, indicating enhanced signal clarity and reduced deviation from expected values. This further underscores the effectiveness of the TSE + PSD integration. Although the TSE + WPD and PSD + WPD integrations exhibited competitive performance, they fell short of surpassing the TSE + PSD integration.

In particular, the triple-modality configuration, which combines TSE, PSD, and WPD, achieved the most substantial performance enhancements, with an RMSE of $\mathbf{2.91 \times 10^{-6}}$, alongside improvements in SAR ($\mathbf{15.6788}$) and NMSE ($\mathbf{0.0649}$). Additionally, MI also reached $\mathbf{3.8388}$, suggesting that the combined modalities offer richer signal representations and information gain. Figure 3 presents radar charts comparing the performance of single-modality, dual-modality, and triple-modality configurations, corresponding to the data shown in Table IV. The charts highlight the performance improvements achieved by combining multiple modalities, with the triple-modality configuration showing the most significant enhancement across all metrics. These findings underscore the potential for exploring additional modality combinations in future research, emphasizing the ability of integrated modalities to capture a more comprehensive representation of underlying signals, leading to improved accuracy and robustness in EEG data processing.

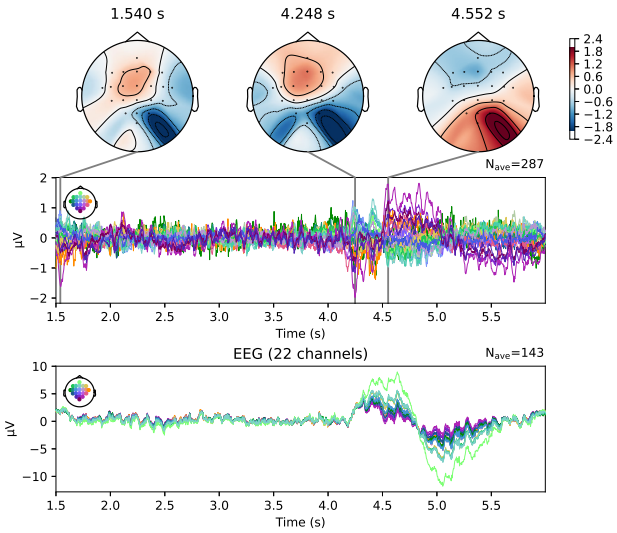
C. Comparison Study

1) Waveform Analysis: Figure 4 illustrates the waveforms and their differences after processing with MASR, alongside the other five comparison methods (Picard-ICA, SSP, OTP, XDAWN, ASR) to the original EEG signal. To ensure methodological consistency, five comparison methods were implemented using the MNE-Python framework, maintaining a standardized processing pipeline [15]. The waveform plots reveal the effects of each method on the signal. Cleaned signals exhibit clear alterations, with notable changes in artifact removal. Each method demonstrates unique cleaning

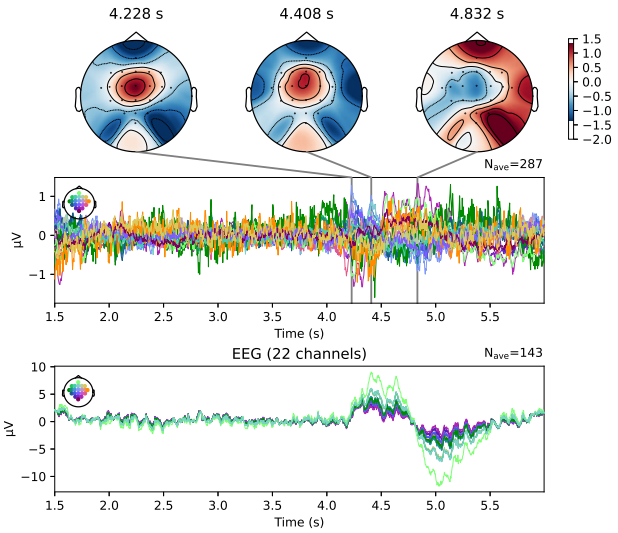
TABLE III
ADAPTIVE GROUPING SETTING ACROSS SUBJECTS

Subject	High-Impact	Medium-Impact	Low-Impact
S1	Fz, FC3, FC1, FC4	FCz, FC2, C5, C3, C1, Cz, C2, C4, C6, CP3, CP1, CPz, CP2, CP4	P1, Pz, P2, POz
S2	Fz, FC3, FC1, FCz, FC4	FC2, C5, C3, C1, Cz, C2, C4, C6, CP3, CP1, CPz, CP2, CP4	P1, Pz, P2, POz
S4	Fz, FCz, FC2, FC4	FC3, FC1, C5, C3, C1, Cz, C2, C4, C6, CP3, CP1, CPz, CP2, CP4	P1, Pz, P2, POz
S5	Fz, FC3, FC4	FCz, FC2, FC1, C5, C3, C1, Cz, C2, C4, C6, CP3, CP1, CPz, CP2, CP4	P1, Pz, P2, POz
S6	Fz, FCz, FC2, Cz	FC3, FC1, FC4, C5, C3, C1, C2, C4, C6, CP1, CPz, CP2, CP4, Pz	CP3, P1, P2, POz
S7	Fz, FC3, FC1, FCz, FC4	FC2, C5, C3, C1, Cz, C2, C4, C6, CP1, CPz, CP2, CP4, Pz	CP3, P1, P2, POz
S8	Fz, FC3, FC2, FC4	FC1, FCz, C5, C3, C1, Cz, C2, C4, C6, CP1, CPz, CP2, CP4, P2	CP3, P1, Pz, POz
S9	Fz, FC3, FC1, FCz, FC2	FC4, C5, C3, C1, Cz, C2, C4, C6, CP1, CPz, CP2	CP3, CP4, P1, Pz, P2, POz

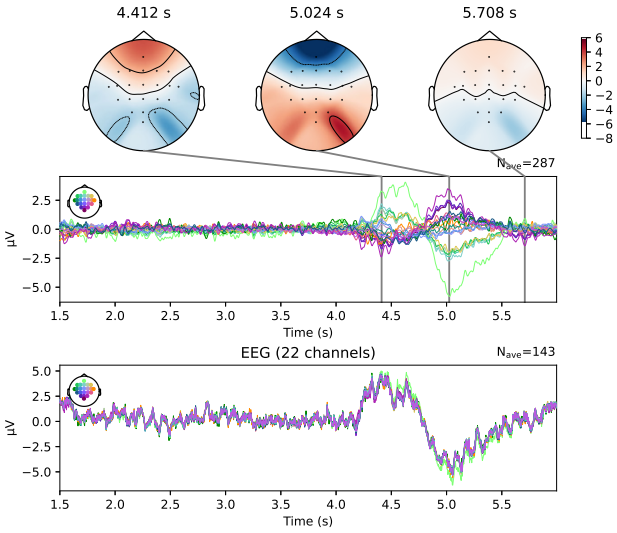
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TETCI.2025.3577504, IEEE Transactions on Emerging Topics in Computational Intelligence



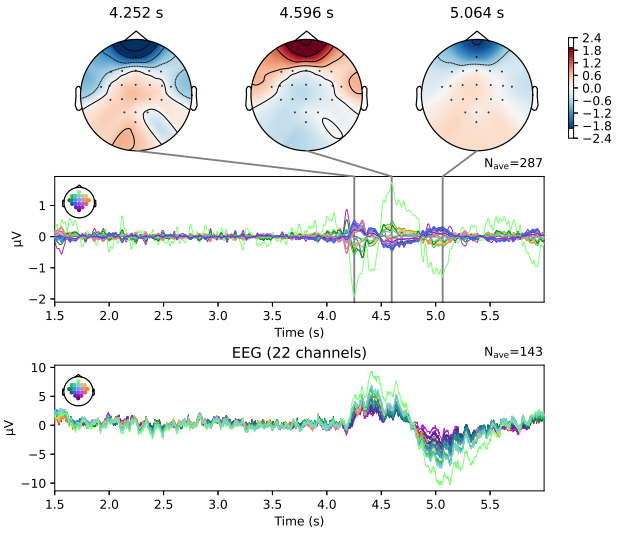
(a) Picard-ICA



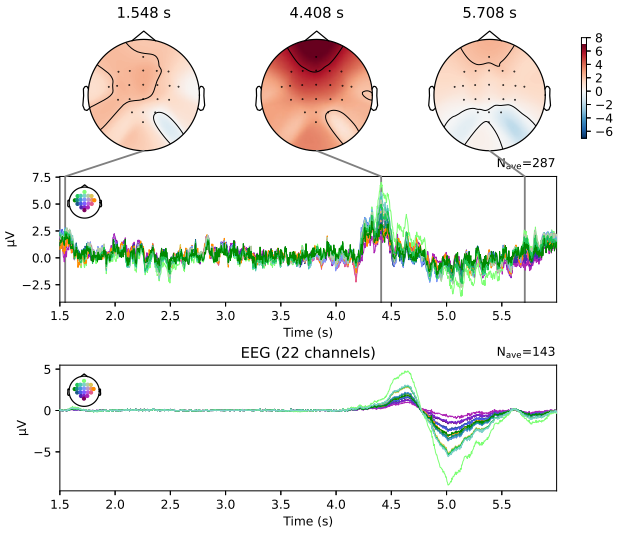
(b) SSP



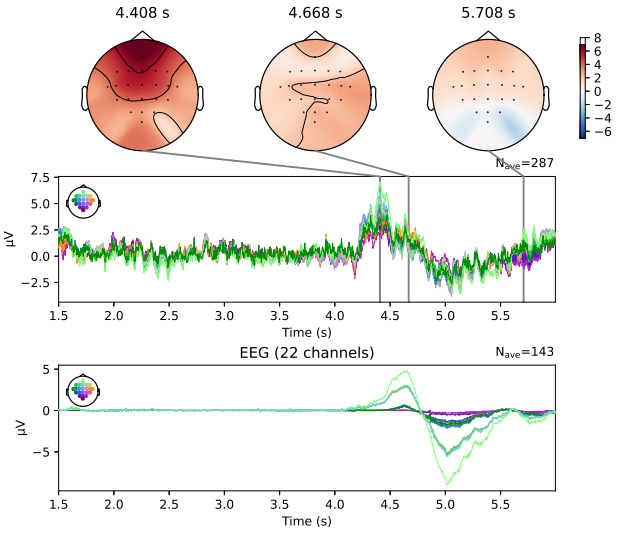
(c) OTP



(d) XDAWN



(e) ASR



(f) MASR

Fig. 4. Comparison of waveforms prior to and following the application of artifact removal methods

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TETCI.2025.3577504, IEEE Transactions on Emerging Topics in Computational Intelligence

characteristics, especially in terms of artifact suppression. Three methods of comparison with MASR are described as follows:

- Picard-ICA is a mathematical technique used to separate a multivariate signal into additive, independent components, particularly useful in identifying and removing artifacts from EEG data by isolating and eliminating noise [1].
- SSP identifies and reduces specific noise components within the signal space [37], suppressing artifacts like those caused by eye blinks by projecting out components recognized as noise.
- OTP is a denoising algorithm that uses a leave-one-out procedure with parallel temporal projection to remove individual sensor noise under the assumption that the sensor array oversamples MEG and EEG fields [20].
- XDAWN is a spatial filtering method designed to improve the signal-to-noise ratio of event-related potential (ERP) responses [31], enhancing target responses by emphasizing them over non-target responses.
- ASR is designed to reduce transient artifacts in EEG data [3], by detecting deviations from a clean EEG signal and reconstructing the signal by removing these disruptions, thereby cleaning the data more effectively.

TABLE IV
ACROSS-SUBJECT RESULTS OF ABLATION EXPERIMENTS FOR DIFFERENT TASK TYPES

Modality Information	RMSE(10^{-6})↓	SAR↑	NMSE↓	MI↑
TSE	3.28	14.2687	0.0686	3.31
PSD	2.95	15.63	0.0663	3.8275
WPD	3.13	14.545	0.0698	3.8063
TSE + PSD	2.93	15.655	0.0661	3.825
TSE + WPD	3.00	15.405	0.0687	3.8188
PSD + WPD	2.95	15.5575	0.066	3.83
TSE + PSD + WPD	**2.91**	**15.6788**	**0.0649**	**3.8388**

The difference plots in Figure 4 show the difference between the cleaned signal and the original signal, reflecting the extent and quality of artifact removal by each method. These visualizations serve as a foundational visual understanding of the performances of methods, laying the groundwork for subsequent quantitative analysis. The topographic maps highlight the changes in the spatial domain after removal of artifacts by further showing the spatial distribution of brain activity at peak time for each method. These maps emphasize how artifact removal influences the representation of EEG signals across scalps.

In the original signals, ocular artifacts appear as slow, high-amplitude fluctuations, predominantly observed between 4 s to 6 s, significantly distorting the temporal features of the waveforms. Following artifact removal, all methods reduced these distortions to varying degrees, with their unique characteristics. Picard-ICA, SSP, OTP, and XDAWN methods occasionally caused signal distortion while suppressing artifacts, whereas ASR and MASR demonstrated superior artifact suppression while preserving signal integrity. Notably, MASR outperformed the methods by retaining continuity and preservation of essential features within the affected time frame, which is vital for the precise extraction of event-related signals in later analyses.

Figure 4 further quantified the deviations between the cleaned signals and the original signal. Picard-ICA, SSP, OTP, and XDAWN primarily suppressed overall amplitude, potentially missing the core regions impacted by ocular artifacts. In contrast, ASR and MASR employed dynamic thresholding to locate and mitigate localized artifact interference. MASR removes artifacts more efficiently and retains more information

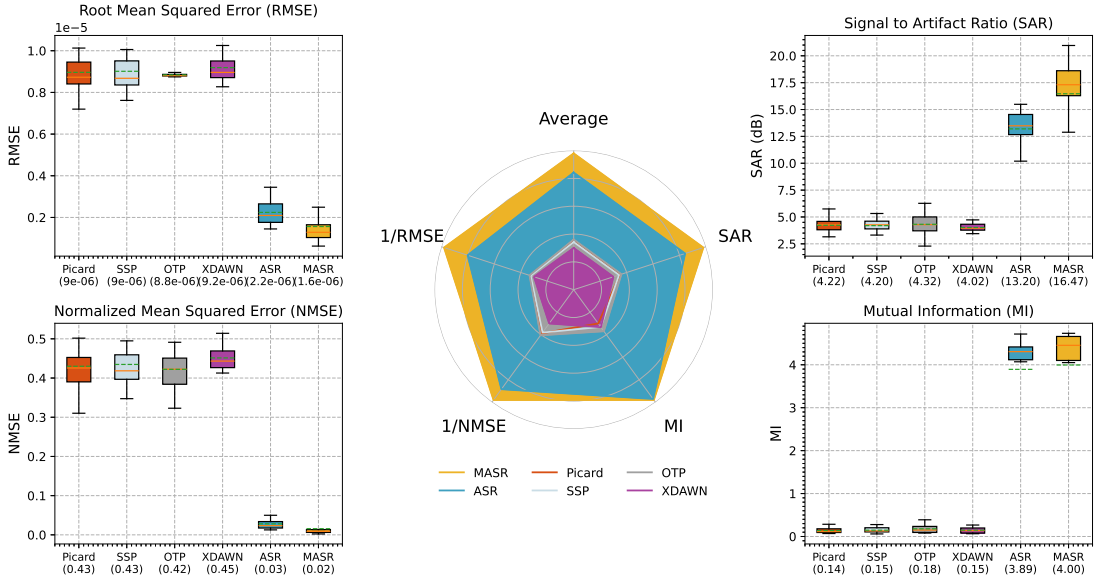


Fig. 5. Comparison of method performance using EEG artifact removal metrics

in less affected channels by targeting the channels that are more affected by artifacts. This advantage is evident in both waveforms and topographical maps, demonstrating the effectiveness of MASR in achieving the primary goals of artifact removal within experimental designs while providing robust EEG processing for complex scenarios.

2) Comparison Based on Metrics: Figure 5 presents a comparison of the six methods using four performance metrics. MASR consistently outperforms Picard-ICA, SSP, OTP, XDAWN, and ASR in minimizing errors and preserving signal fidelity. MASR's multimodal framework integrates multiple artifact sources, enabling effective noise reduction while maintaining the integrity of the original EEG signal.

In contrast to ASR, which relies on single-modality artifact removal, MASR utilizes adaptive thresholding across modalities, ensuring precise differentiation between noise and signal. This adaptive thresholding enhances accuracy and stability, as evidenced by MASR's superior performance in metrics such as RMSE, SAR, and MI. MASR achieves a balance between artifact suppression and signal preservation, making it especially effective in high-noise contexts, unlike other traditional methods, which employ rigid and less flexible methodologies.

The multimodal design of MASR provides an additional advantage by leveraging complementary information from various EEG features, enabling better cross-validation of noise components and retention of meaningful signal structures, particularly evident in the MI results. The comparison study illustrates MASR's comprehensive advantages across all performance metrics, including the enhancement of signal-to-artifact ratio, minimization of reconstruction errors, and preservation of critical signal features.

IV. Conclusion

This study evaluated the effectiveness of the Multimodal Artifact Subspace Reconstruction (MASR) method for EEG artifact removal using the BCI Competition IV-2a dataset. In contrast to traditional methods like ICA, SSP, and ASR, which use fixed thresholds, MASR employs a data-driven approach that dynamically determines the threshold using principal component analysis. This allows for more precise artifact removal by grouping channels based on statistical criteria without relying on user-defined thresholds. Comparative evaluation using RMSE, SAR, NMSE, and MI showed that MASR outperformed the other methods in error minimization and signal preservation. Ablation experiments confirmed that combining TSE and PSD provided strong performance, and adding the third modality further improved accuracy. Future work could involve tuning MASR for different EEG signal types, testing various TSE, PSD, and WPD configurations, and incorporating non-linear dependency measures to better capture complex interactions between artifacts and EEG signals. Additionally, optimizing the computational efficiency of MASR could improve its applicability in real-time EEG processing.

References

[1] P. Ablin, J. -F. Cardoso and A. Gramfort, "Faster independent component analysis by preconditioning with Hessian approximations," *IEEE Transactions on Signal Processing*, vol. 66, no. 15, pp. 4040-4049, 2018.

[2] J. N. Acosta, "Multimodal biomedical AI," *Nature Medicine*, vol. 28, 2022.

[3] S. Blum, N. S. J. Jacobsen, M. G. Bleichner, and S. Debener, "A Riemannian modification of artifact subspace reconstruction for EEG artifact handling," *Frontiers in Human Neuroscience*, vol. 13, p. 141, 2019.

[4] J. Cao et al., "Brain functional and effective connectivity based on electroencephalography recordings: A review," *Human Brain Mapping*, vol. 43, no. 2, pp. 860–879, 2022.

[5] J. Cao, Y. Fang, X. Cui, R. Zheng, T. Jiang and F. Gao, "Synchronized video and EEG based childhood epilepsy seizure detection," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 8, no. 6, pp. 3742-3753, 2024.

[6] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D: Nonlinear Phenomena*, vol. 110, no. 1–2, pp. 43–50, 1997.

[7] A. Chaddad, Y. Wu, R. Kateb, and A. Bouridane, "Electroencephalography signal processing: A comprehensive review and analysis of methods and techniques," *Sensors*, vol. 23, no. 14, p. 6434, 2023.

[8] C.-Y. Chang, S.-H. Hsu, L. Pion-Tonachini, and T.-P. Jung, "Evaluation of artifact subspace reconstruction for automatic artifact components removal in multi-channel EEG recordings," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 4, pp. 1114–1121, 2020.

[9] X. Chen, Y. Zhang, S. Cheng, and P. Xie, "Transfer spectral entropy and application to functional corticomuscular coupling," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 5, 2019.

[10] Z. Chen, R. Yang, M. Huang, Z. Wang and X. Liu, "Electrode domain adaptation network: Minimizing the difference across electrodes in single-source to single-target motor imagery classification," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 8, no. 2, pp. 1994-2008, 2024.

[11] F. Cui, R. Wang, W. Ding, Y. Chen, and L. Huang, "A novel DE-CNN-BiLSTM multi-fusion model for EEG emotion recognition," *Mathematics*, vol. 10, no. 4, p. 582, 2022.

[12] J. Escudero, R. Hornero, D. Abásolo, and A. Fernández, "Quantitative evaluation of artifact removal in real magnetoencephalogram signals with blind source separation," *Annals of Biomedical Engineering*, vol. 39, no. 8, pp. 2274–2286, 2011.

[13] X. Gao, F. Deng, W. Shang, X. Zhao and S. Li, "Attack-resilient asynchronous state estimation of interval type-2 fuzzy systems under stochastic protocols," *International Journal of Systems Science*, vol. 55, no. 13, pp. 2688-2700, 2024.

[14] S. K. Goh, H. A. Abbass, K. C. Tan, A. Al-Mamun, C. Wang, and C. Guan, "Automatic EEG artifact removal techniques by detecting influential independent components," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 1, no. 4, pp. 270–279, 2017.

[15] A. Gramfort et al., "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, 2013.

[16] X. Jiang, G.-B. Bian, and Z. Tian, "Removal of artifacts from EEG signals: A review," *Sensors*, vol. 19, no. 5, p. 987, 2019.

[17] Y. Jiang, Y. Zhang, C. Lin, D. Wu, and C.-T. Lin, "EEG-based driver drowsiness estimation using an online multi-view and transfer TSK fuzzy system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1752–1764, 2021.

[18] G. Karakaya, S. Galelli, S. D. Ahipasaoglu, and R. Taormina, "Identifying (Quasi) equally informative subsets in feature selection problems for classification: A max-relevance min-redundancy approach," *IEEE Transactions on Cybernetics*, vol. 46, no. 6, pp. 1424–1437, 2016.

[19] Y. Kutlu and D. Kuntalp, "Feature extraction for ECG heartbeats using higher order statistics of WPD coefficients," *Computer Methods and Programs in Biomedicine*, vol. 105, no. 3, pp. 257–267, 2012.

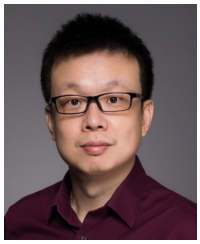
[20] E. Larson and S. Taulu, "Reducing Sensor Noise in MEG and EEG Recordings Using Oversampled Temporal Projection," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 5, pp. 1002–1013, 2018.

[21] J. Lopez-Calderon and S. J. Luck, "ERPLAB: an open-source toolbox for the analysis of event-related potentials," *Frontiers in Human Neuroscience*, vol. 8, 2014.

[22] M. Li, K. Zhang, Y. Ma and B. Jiang, "Prescribed-time fault-tolerant control for the formation of quadrotors based on fully-actuated system approaches," *International Journal of Systems Science*, vol. 55, no. 12, pp. 2541-2555, 2024.

[23] C. Ma, P. Cheng and C. Cai, "Localization and mapping method based on multimodal information fusion and deep learning for dynamic object removal," *International Journal of Network Dynamics and Intelligence*, vol. 3, no. 2, art. no. 100008, 2024.
[24] N. Mammone, F. La Foresta, and F. C. Morabito, "Automatic artifact rejection from multichannel scalp EEG by wavelet ICA," *IEEE Sensors Journal*, vol. 12, no. 3, pp. 533–542, 2012.
[25] M. M. N. Mannan, M. A. Kamran, and M. Y. Jeong, "Identification and removal of physiological artifacts from electroencephalogram signals: A review," *IEEE Access*, vol. 6, pp. 30630–30652, 2018.
[26] W. Mumtaz, S. Rasheed, and A. Irfan, "Review of challenges associated with the EEG artifact removal methods," *Biomedical Signal Processing and Control*, vol. 68, p. 102741, 2021.
[27] J. D. Power, M. Plitt, T. O. Laumann, and A. Martin, "Sources and implications of whole-brain fMRI signals in humans," *NeuroImage*, vol. 146, pp. 609–625, 2017.
[28] X. Qian and B. Cui, "A mobile sensing approach to distributed consensus filtering of 2D stochastic nonlinear parabolic systems with disturbances," *Systems Science & Control Engineering*, vol. 11, no. 1, art. no. 2167885, 2023.
[29] D. Ramachandram and G. W. Taylor, "Deep multimodal learning: A survey on recent advances and trends," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 96–108, 2017.
[30] F. Riaz, A. Hassan, S. Rehman, I. K. Niazi, and K. Dremstrup, "EMD-based temporal and spectral features for the classification of EEG signals using supervised learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 28–35, 2016.
[31] B. Rivet, A. Souloumiac, V. Attina, and G. Gibert, "xDAWN Algorithm to Enhance Evoked Potentials: Application to Brain–Computer Interface," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 8, pp. 2035–2043, 2009.
[32] M. T. Sadiq et al., "Motor imagery BCI classification based on multivariate variational mode decomposition," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, no. 5, pp. 1177-1189, 2022.
[33] N. J. Sairamya, M. S. P. Subathra, E. S. Suviseshamuthu, and S. Thomas George, "A new approach for automatic detection of focal EEG signals using wavelet packet decomposition and quad binary pattern method," *Biomedical Signal Processing and Control*, vol. 63, p. 102096, 2021.
[34] K. Tan, W. Huang, X. Liu, J. Hu, and S. Dong, "A multi-modal fusion framework based on multi-task correlation learning for cancer prognosis prediction," *Artificial Intelligence in Medicine*, vol. 126, p. 102260, 2022.
[35] A. Tharwat, "Independent component analysis: An introduction," *Applied Computing and Informatics*, vol. 17, no. 2, pp. 222–249, 2021.
[36] M. Tangermann et al., "Review of the BCI Competition IV," *Frontiers in Neuroscience*, vol. 6, 2012.
[37] M. A. Uusitalo and R. J. Ilmoniemi, "Signal-space projection method for separating MEG or EEG into components," *Medical & Biological Engineering & Computing*, vol. 35, no. 2, pp. 135–140, 1997.
[38] F.-B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki, "Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives," *Progress in Neurobiology*, vol. 90, no. 4, pp. 418–438, 2010.
[39] J. Wang, L. Shi, W. Wang and Z. -G. Hou, "Efficient brain decoding based on adaptive EEG channel selection and transformation," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, no. 6, pp. 1314-1323, 2022.
[40] M. Wang, X. Cui, T. Wang, T. Jiang, F. Gao, and J. Cao, "Eye blink artifact detection based on multi-dimensional EEG feature fusion and optimization," *Biomedical Signal Processing and Control*, vol. 83, p. 104657, 2023.
[41] M. Wei, M. Huang and J. Ni, "Cross-subject EEG channel selection method for lower limb brain-computer interface," *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 3, art. no. 100008, 2023.
[42] D. Wu, Y. Xu, and B.-L. Lu, "Transfer learning for EEG-based brain–computer interfaces: A review of progress made since 2016," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 1, pp. 4–19, 2022.
[43] D. Xie, C. Deng, C. Li, X. Liu, and D. Tao, "Multi-task consistency-preserving adversarial hashing for cross-modal retrieval," *IEEE Transactions on Image Processing*, vol. 29, pp. 3626–3637, 2020.
[44] J. Zhang, Z. Yin, P. Chen, and S. Nichele, "Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review," *Information Fusion*, vol. 59, pp. 103–126, 2020.
[45] R. Zhang, H. Liu, Y. Liu and H. Tan, "Dynamic event-triggered state estimation for discrete-time delayed switched neural networks with constrained bit rate," *Systems Science & Control Engineering*, vol. 12, no. 1, art. no. 2334304, 2024.
[46] Y. Zhang et al., "Multi-modality fusion & inductive knowledge transfer underlying non-sparse multi-kernel learning and distribution adaption," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 20, no. 4, pp. 2387–2397, 2023.
[47] L. Zheng et al., "A power spectrum pattern difference-based time-frequency sub-band selection Method for MI-EEG classification," *IEEE Sensors Journal*, vol. 22, no. 12, pp. 11928–11939, 2022.

Wenlong You received the B.Eng. degree from the University of Jinan, China, in 2022 and the M.Res. degree in Computer Science from the University of Liverpool, UK, in 2025. He is currently working as a Research Assistant at The Hong Kong Polytechnic University. His research interests include multimodal learning, signal processing, and brain–computer interfaces.



Rui Yang (Senior Member, IEEE) received the B.Eng. degree in Computer Engineering and the Ph.D. degree in Electrical and Computer Engineering from National University of Singapore in 2008 and 2013 respectively. He is currently an Associate Professor in the School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, China, and an Honorary Lecturer in the Department of Computer Science, University of Liverpool, Liverpool, United Kingdom. His research interests include machine learning based data analysis and applications. He is the author or co-author of several technical papers and also a very active reviewer for many international journals and conferences. Dr. Yang is currently serving as an Associate Editor for Neurocomputing, Cognitive Computation, and IEEE Transactions on Instrumentation and Measurement.

Chengxuan Qin received the MRes degree in Pattern Recognition and Intelligent Systems from the University of Liverpool in 2023. He is pursuing a Ph.D. degree at the University of Liverpool, China. His research interests include brain-computer interfaces, temporal signal analysis, and machine learning.

Mengjie Huang received the Ph.D. degree from National University of Singapore in 2014, and the B.Eng degree from Sichuan University in 2009. She is now an Associate Professor in the Design School, Xi'an Jiaotong-Liverpool University, Suzhou, China. Dr Huang's current research interests include human-computer interaction and applications.

Zidong Wang (Fellow, IEEE) was born in Jiangsu, China, in 1966. He received the B.Sc. degree in mathematics from Suzhou University, Suzhou, China, in 1986, and the M.Sc. degree in applied mathematics and the Ph.D. degree in electrical engineering from the Nanjing University of Science and Technology, Nanjing, China, in 1990 and 1994, respectively.

From 1990 to 2002, he held teaching and research appointments in universities in China, Germany, and the U.K. He is currently a Professor of dynamical systems and computing with the Department of Computer Science, Brunel University London, Uxbridge, U.K. He has published more than 600 articles in international journals. His research interests include dynamical systems, signal processing, bioinformatics, control theory, and applications.

Prof. Wang is a member of the Academia Europaea and the European Academy of Sciences and Arts, an Academician of the International Academy for Systems and Cybernetic Sciences, a fellow of the Royal Statistical Society, and a member of the Program Committee of many international conferences. He holds the Alexander von Humboldt Research Fellowship of Germany, the JSPS Research Fellowship of Japan, and the William Mong Visiting Research Fellowship of Hong Kong. He serves (or has served) as the Editor-in-Chief for International Journal of Systems Science, Neurocomputing, and Systems Science and Control Engineering; and an Associate Editor for 12 international journals, including IEEE Transactions on Automatic Control, IEEE Transactions on Control Systems Technology, IEEE Transactions on Neural Networks and Learning Systems, IEEE Transactions on Signal Processing, and IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews.