

# State-sensitive convolutional sparse coding for potential biomarker identification in brain signals

Puli WANG<sup>1,4</sup>, Yu QI<sup>2,3,4\*</sup> & Gang PAN<sup>1,4\*</sup><sup>1</sup>College of Computer Science and Technology, Zhejiang University, Hangzhou 310000, China;<sup>2</sup>MOE Frontier Science Center for Brain Science and Brain-machine Integration, Zhejiang University, Hangzhou 310000, China;<sup>3</sup>Affiliated Mental Health Center & Hangzhou Seventh People's Hospital, Zhejiang University School of Medicine, Hangzhou 310000, China;<sup>4</sup>State Key Laboratory of Brain-Machine Intelligence, Zhejiang University, Hangzhou 310000, China

Received 9 May 2023/Revised 24 July 2023/Accepted 28 September 2023/Published online 8 April 2024

**Abstract** The identification of prototypical waveforms, such as sleep spindles and epileptic spikes, is crucial for the diagnosis of neurological disorders. These prototypical waveforms are usually recurrently presented in certain brain states, serving as potential biomarkers for clinical evaluations. Convolutional sparse coding (CSC) approaches have demonstrated strength in identifying recurrent patterns in time-series. However, existing CSC approaches do not explicitly explore state-specific patterns, making it difficult to identify state-related biomarkers. To address this problem, we propose state-sensitive CSC to learn state-specific prototypical waveforms. Specifically, we model signals of a certain state with specific waveforms that only appear frequently in this state and background waveforms that are independent of states. Based on this, state-sensitive CSC separates state-specific waveforms from background ones explicitly by incorporating incoherence constraints into optimizations. Experiments with epilepsy brain signals demonstrate that our approach can effectively identify prototypical waveforms in pre-ictal states, providing potential biomarkers for seizure prediction. Our approach provides a promising tool for automatic biomarker candidate identification.

**Keywords** convolutional sparse coding, prototypical waveforms, state-specific dictionary, biomarker identification, neural signal processing

## 1 Introduction

In brain signals, prototypical waveforms can provide meaningful signatures for identifying complex brain cognitive processes [1–3]. For instance, the alternations of sleep spindles, which are spontaneous 12 to 14 Hz transient oscillations, indicate several disorders such as schizophrenia [4,5], autism [6,7], and mental retardation [8]. Epileptic spikes [9–11] and high-frequency oscillations (HFO) [12] such as ripples [13] are highly relevant to epileptic onsets and can assist with seizure prediction and diagnosis. Therefore, the discovery and identification of these prototypical waveforms aid in detecting prognostic biomarkers for neurological disorders.

To identify these prototypical waveforms, efforts have been made, including knowledge-based approaches and data-driven approaches. Traditional knowledge-based approaches usually assume the characteristics of the prototypical waveforms in advance, such as frequency ranges [14] and temporal patterns. For instance, the most widely used knowledge-based approaches define prototypical waveforms as the high power activity in time-frequency representations that exceeds a pre-set power threshold and falls within a predefined frequency range [15,16]. While more advanced knowledge-based approaches have been proposed to improve sensitivity to the high power activity by accounting for aperiodic background activity [17,18], these methods still rely on several predefined assumptions. Data-driven approaches aim to reduce the dependence on predefined assumptions by learning prototypical waveforms directly from neural signals. Typical data-driven approaches include empirical mode decomposition (EMD) [19,20], cycle-by-cycle analysis [21], and brief amplitude undulation detection (BAU) [22–24], which automatically identify

\* Corresponding author (email: qiyu@zju.edu.cn, gpan@zju.edu.cn)

prototypical waveforms based on their temporal shapes. As for learning waveforms across channels, methods based on independent component analysis (ICA) [25], hidden Markov modeling (HMM) [26, 27], and electroencephalogram (EEG) microstates [28, 29] have been utilized to discover waveforms with certain characteristics in spatial patterns.

Most recently, dictionary learning approaches [30] have demonstrated superior performance in learning prototypical waveforms. Convolutional sparse coding (CSC) is a specific type of convolutional dictionary learning [31–33]. The core idea of CSC is to decompose neural signals into combinations of recurrent shift-invariant patterns as prototypical waveforms, which are activated by the typical neural activity [34] or featureless background activity that maintains basic functionality [35]. CSC-based approaches have achieved promising performance in detecting and characterizing prototypical waveforms related to aging trends [36], auditory and visual stimuli [37].

One limitation of existing CSC-based approaches lies in that, they mostly do not explicitly consider the specific characteristics of learned waveforms. On the other hand, state-specificity is a notable characteristic of prognostic biomarkers, which implies their frequent appearance only in a certain state. Thus, it is difficult for existing CSC-based approaches to be directly applied to the biomarker identification due to their unsatisfactory performance of distinguishing state-specific biomarker candidates from irrelevant background waveforms. This motivates us to develop novel CSC-based approaches that incorporate state-specificity explicitly into the process of learning prototypical waveforms, which can provide candidates for prognostic biomarkers.

To this end, we propose a novel data-driven approach based on CSC, called state-sensitive CSC, for the automatic identification of state-specific prototypical waveforms. We first propose to model signals of a particular state with state-specific waveforms that only appear frequently in this state and background waveforms that are independent of states. Based on this, we further introduce a novel algorithm to learn state-specific prototypical waveforms, where the separation of state-specific waveforms and background waveforms is guaranteed explicitly by regularization terms on their incoherence. Quantitative and analytical results on epilepsy brain signals demonstrate that our approach can provide novel potential biomarkers for predicting seizure onsets. The main contributions of this paper are summarized as follows.

- (1) We propose to model signals with a combination of state-specific waveforms that repetitively appear only in a certain state and background waveforms that share among multiple states.
- (2) We introduce a novel state-sensitive CSC algorithm to learn state-specific prototypical waveforms, which can serve as a generalizable framework for automatic biomarker candidate identification.
- (3) Experiments with epilepsy brain signals demonstrate that our approach can successfully identify state-specific prototypical waveforms in pre-ictal states, which provide new features for seizure prediction.

## 2 Related work

### 2.1 Sparse coding for discriminative dictionary learning

Sparse coding (SC) models signal as a linear combination of waveforms in an overcomplete dictionary. For the sake of more discriminative representations, two types of SC-based approaches are proposed, where constraints on sparse codes or waveforms are incorporated into the learning process.

**One type of SC-based approaches exploit constraints on sparse codes for discriminability.** For instance, within-group sparse and group-wise sparse constraints [38] are utilized for learning class-specific and subject-specific waveforms, which effectively reduce the required training samples from the target subject. Similarly, tree-guided group sparse learning is proposed to select waveforms that are specific to chosen brain regions with the region spatial constraint on sparse codes. These selected waveforms are further demonstrated to be potential clinical biomarkers for resting-state fMRI-based schizophrenia diagnosis [39]. Besides the application in brain signal processing, discriminative dictionary learning and classifier training based on the sparse codes are unified for face and object recognition [40–42].

**The other type of SC-based approaches exploit constraints on waveforms for discriminability.** For example, shared and subject-specific dictionary learning (ShSSDL) is proposed to extract shared and subject-specific waveforms from multisubject fMRI data using the incoherence constraints on waveforms [43]. Likewise, images are decomposed into shared and category-specific dictionaries, where category-specific dictionaries are learned under the incoherence regularization terms, to amplify their originally subtle differences for fine-grained categorization [44]. Moreover, data-driven elements are opti-

mized with task stimulus curves fixed as predefined elements in the dictionary learning procedure, which can be applied to inferring discriminative task-related functional networks from fMRI data [45].

Despite the superior performance of learning discriminative dictionaries, these SC-based approaches do not fully capture the temporal dynamics of prototypical waveforms, which can be important for identifying prognostic biomarkers that are only present in certain states. Furthermore, waveforms learned by SC are often duplicated with a time-shift, leading to the redundancy of resultant biomarkers.

## 2.2 CSC

CSC is widely-used in several areas including audio signal processing [46], biomedical imaging [47], and computer vision [48, 49]. When applied to brain signal processing, CSC represents these prototypical waveforms with recurrent shift-invariant patterns, which are activated by independent sparse sources. For instance, recurrent waveforms within EEG are learned via single-channel ICA with the post-hoc waveform selection [50]. Alpha-stable CSC is developed to extract meaningful waveforms from signals contaminated by severe non-Gaussian noises [32]. As for the multi-channel learning, multivariate CSC is introduced to detect simultaneous waveforms under the rank-1 constraint on sparse sources [33].

To distinguish discriminative waveforms from background ones, additional selections are conducted after the standard CSC-based dictionary learning. For example, discriminative waveforms are identified as snippets whose average correntropy-based similarity with others is smaller [51]. This approach is based on the observation that the amplitude distribution of neuromodulations deviates significantly from Gaussian, while that of background activity follows a Gaussian distribution. Furthermore, in order to identify task-related waveforms, post-hoc selections based on driven point processes (DriPP) [37] and rate of occurrence [36] measure the correlation between waveforms and specific stimuli quantitatively.

Though existing CSC-based approaches can capture temporal dynamics of shift-invariant waveforms compared with SC-based approaches, their approaches of selecting waveforms are not well-suited for biomarker identification. Thus, inspired by the success of discriminative dictionary learning via SC, we propose a novel CSC-based approach, which incorporates state-specificity explicitly into the learning process.

## 3 State-sensitive CSC

### 3.1 Existing brain signal model with CSC

The sparse and repeating occurrence of prototypical waveforms has been deemed to be the feature of brain dynamics. These ordered patterns bring transient organization to complex and chaotic brain dynamical systems [52]. In order to learn these recurrent patterns, CSC decomposes brain signals into a mixture of waveforms, which are activated by independent sparse sources [31, 50, 53].

According to (1), the single-channel and band-passed brain signal,  $x(t)$ , is assumed as the linear summation of components reconstructed by waveforms  $\hat{x}(t)$  and additive Gaussian noises  $\epsilon(t)$ . The reconstructed signal  $\hat{x}(t)$  is further formulated by sparse impulse responses of distinct waveforms based on (2).

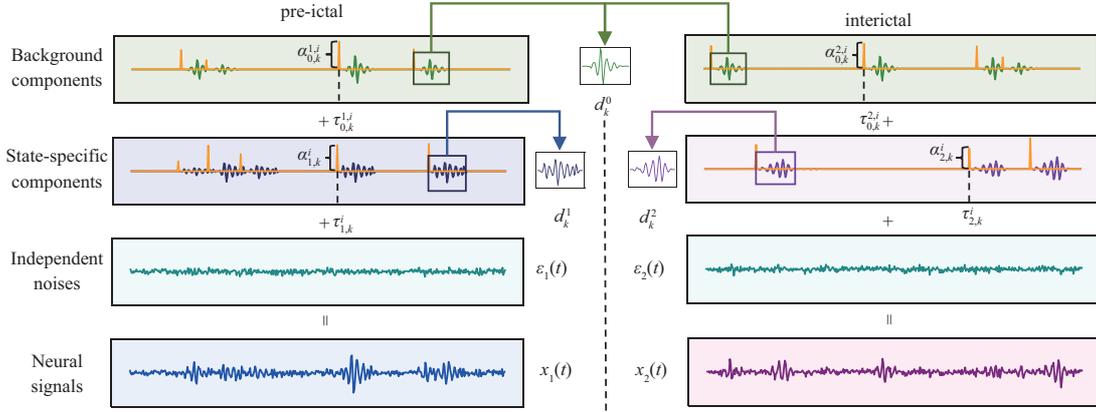
$$x(t) = \hat{x}(t) + \epsilon(t), \quad (1)$$

$$\hat{x}(t) = \sum_{k=1}^w \sum_{i=1}^{n_k} a_k^i d_k(t - \tau_k^i) \quad (\|d_k\|_2 = 1, \forall k). \quad (2)$$

Here,  $\{d_k\}_{k=1}^w$  is a dictionary composed of  $w$  data-driven waveforms with the unit  $\ell_2$ -norm. The  $i$ -th activation of  $d_k$  is expressed as a phasic event, which is depicted by the projection coefficient  $a_k^i$  and the timing  $\tau_k^i$ .  $n_k$  denotes the number of phasic events that evoke  $d_k$ . It is worth noting that  $d_k$  defined by CSC is irrelevant to the state of  $x(t)$ .

### 3.2 Extended brain signal model with state-specific waveforms

Though the model presented in (1) and (2) can extract various prototypical waveforms, differentiating state-specific waveforms from those generated by background activity remains a challenge for these methods. This difficulty arises because the decomposition is state-independent, resulting in the waveform overlap across different states. This overlap may obscure the unique characteristics of a particular state,



**Figure 1** (Color online) Illustration of our extended brain signal model with state-specific waveforms for the task of seizure prediction.

thereby limiting the use of these waveforms as biomarkers. As a result, this motivates us to develop novel brain modeling approaches that can account for state-specificity.

To this end, we propose a novel brain model that considers the presence of two different types of waveforms in signals from a particular state. The first type of waveform is state-specific and occurs recurrently in a certain state, making it a candidate for a biomarker. The second type is generated by featureless background activity and is present in all states. By encoding neural signals with these state-specific waveforms, we can enhance the differences between states and obtain more discriminative representations.

Mathematically, we denote the single-channel and band-passed signal from a particular state among the total  $C$  states as  $x_c(t)$ , where  $c \in \{1, \dots, C\}$ . Then, according to (3),  $x_c(t)$  can be expressed as a summation of waveform components  $\hat{x}_c(t)$  and independent Gaussian noise  $\epsilon_c(t)$ .  $\hat{x}_c(t)$  is further formulated as the sparse activations of state-specific and background waveforms:

$$x_c(t) = \hat{x}_c(t) + \epsilon_c(t), \quad (3)$$

$$\hat{x}_c(t) = \sum_{k=1}^{w_0} \sum_{i=1}^{n_k^{0,c}} a_{0,k}^{c,i} d_k^0(t - \tau_{0,k}^{c,i}) + \sum_{k=1}^{w_c} \sum_{i=1}^{n_k^c} a_{c,k}^i d_k^c(t - \tau_{c,k}^i) \quad (\|d_k^0\|_2 = 1, \forall k; \|d_k^c\|_2 = 1, \forall k), \quad (4)$$

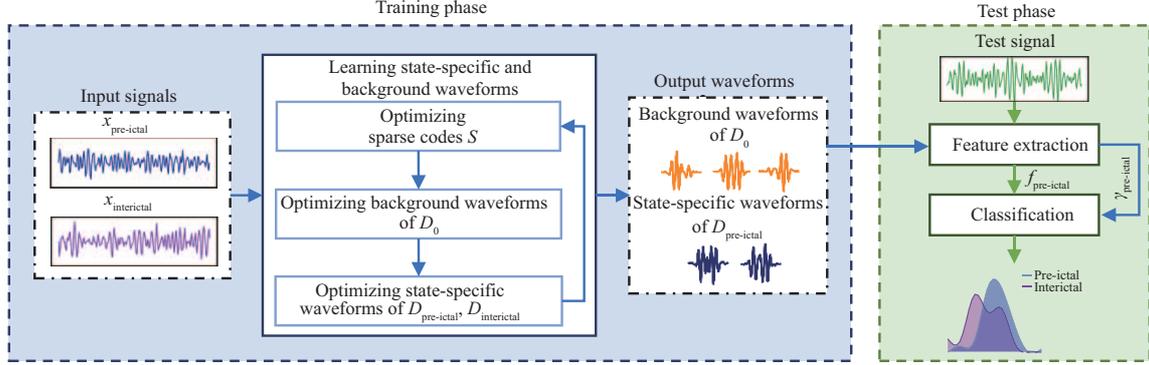
where the background ( $D_0$ ) and state-specific ( $D_c$ ) dictionaries consist of  $w_0$  and  $w_c$  waveforms with the unit  $\ell_2$ -norm, respectively:  $D_0 = \{d_k^0\}_{k=1}^{w_0}$  and  $D_c = \{d_k^c\}_{k=1}^{w_c}$ . The  $i$ -th activation of a background waveform  $d_k^0$  is defined by a non-zero amplitude  $a_{0,k}^{c,i}$  and timing  $\tau_{0,k}^{c,i}$ . Similarly, the  $i$ -th event evoked by a state-specific waveform  $d_k^c$  is represented by a non-zero amplitude  $a_{c,k}^i$  and timing  $\tau_{c,k}^i$ . The total number of phasic events evoked by  $d_k^0$  and  $d_k^c$  are denoted as  $n_k^{0,c}$  and  $n_k^c$ , respectively.

For convenience, we represent the sparse codes of  $D_0$  and  $D_c$  as sets containing pairs of coefficients denoting amplitudes and timings:  $S_0^c = \{(a_{0,k}^{c,i}, \tau_{0,k}^{c,i})\}_{i=1, k=1}^{n_k^{0,c}, w_0}$ ,  $S_c = \{(a_{c,k}^i, \tau_{c,k}^i)\}_{i=1, k=1}^{n_k^c, w_c}$ .  $D$  and  $S$  consist of all waveforms and their corresponding encoding of signals:  $D = D_0 \cup \{D_c\}_{c=1}^C$ ,  $S = \{S_0^c, S_c\}_{c=1}^C$ . It is worth noting that the state-specific waveform  $d_k^c$  defined in (4) is relevant to the state  $c$ . An example of the model is illustrated in Figure 1 for the task of seizure prediction.

### 3.3 Learning state-specific waveforms

To obtain state-specific and background waveforms, we propose an optimization algorithm based on the model defined in (3) and (4). The main objective of this algorithm is to ensure the state-specificity of waveforms during the learning process. Therefore, in addition to minimizing the  $\ell_2$ -norm error of  $D_0$  and  $D_c$  to reconstruct  $x_c$  and enforcing sparsity of  $S$  as in standard CSC, we also introduce constraints on the self-incoherence and cross-incoherence of waveforms to ensure state-specificity [54].

To measure the similarity between two waveforms  $d_i$  and  $d_j$ , we define a function  $\text{corr}$  that considers any time-shift via convolution:  $\text{corr}(d_i, d_j) = \max(d_i * \tilde{d}_j)$ , where  $*$  denotes the discrete convolution operator and  $\tilde{d}_j$  is the time-reversed version of  $d_j$  with a length of  $L$  time points. We further define



**Figure 2** (Color online) Framework of our method to learn state-specific waveforms and validate their performance as potential prognostic biomarkers under the context of seizure prediction.

$\text{corr}(D_i, D_j)$  to measure the similarity between all pairs of waveforms selected from  $D_i$  and  $D_j$ . For convenience, we denote  $D_{-c}$  as a set containing all dictionaries except  $D_c$ .

We then add a regularization term  $\sum_{c=0}^C \text{corr}(D_c, D_{-c})$  to enforce the cross-incoherence of  $D_c$ . Without this constraint, the sparse codes of the background dictionary may be all zeros since signals can be best reconstructed merely by the state-specific dictionary which contains background waveforms. Moreover, we also add the term  $\sum_{c=0}^C \text{corr}(D_c, D_c)$  to enforce the self-incoherence within  $D_c$ , which prevents duplicated waveforms with time-shifts [55] and stabilizes the resultant waveforms [56]. Therefore, the design problem can be posed as a least-square optimization with the additional constraints described above:

$$\begin{aligned} \min_{D, S} & \left\{ \sum_{c=1}^C \|x_c(t) - g(S_0^c, D_0, t) - g(S_c, D_c, t)\|_F \right. \\ & \left. + \lambda_w \sum_{c=1}^C (|S_0^c| + |S_c|) + \lambda_c \sum_{c=0}^C \text{corr}(D_c, D_{-c}) + \lambda_s \sum_{c=0}^C \text{corr}(D_c, D_c) \right\} \\ \text{s.t.} & \|d_k^0\|_2 = 1, \forall k; \|d_k^c\|_2 = 1, \forall c, k. \end{aligned} \quad (5)$$

In this equation,  $\|\cdot\|_F$  denotes the Frobenius norm. The term  $|S|$  indicates the total number of coefficient pairs in the sparse code  $S$ , and the expression  $\sum_{c=1}^C (|S_0^c| + |S_c|)$  represents the sparsity of waveforms. The function  $g$  computes the reconstructed signal comprising input waveforms and their corresponding sparse codes as follows:  $g(S_0^c, D_0, t) = \sum_{k=1}^{w_0} \sum_{i=1}^{n_k^{0,c}} a_{0,k}^{c,i} d_k^0(t - \tau_{0,k}^{c,i})$ ,  $g(S_c, D_c, t) = \sum_{k=1}^{w_c} \sum_{i=1}^{n_k^c} a_{c,k}^i d_k^c(t - \tau_{c,k}^i)$ .  $\lambda_w$ ,  $\lambda_c$ , and  $\lambda_s$  are constant weights of the sparsity, cross-incoherence, and self-coherence constraints, respectively. The unit  $\ell_2$ -norm constraint on each element of  $D$  is used to avoid the trivial solution.

The objective function is not jointly convex over  $S$  and  $D$ . However, we can optimize sparse codes, background and state-specific waveforms alternately while fixing resting variables, following strategies mentioned in [31, 44, 53]. To ensure state-specificity, we propose to extract background and state-specific waveforms from different subsets of signals under the constraint of proposed regularization terms. Each optimizing procedure is explained in detail in this section, and the general framework is depicted in Figure 2.

### 3.3.1 Optimizing sparse codes $S$

To solve for the optimal sparse codes  $S$ , we fix all dictionaries  $D$  and rewrite the objective function formulated in (5) as shown below:

$$\min_S \left\{ \sum_{c=1}^C \|x_c(t) - g(S_0^c, D_0, t) - g(S_c, D_c, t)\|_F + \lambda_w \sum_{c=1}^C (|S_0^c| + |S_c|) \right\}. \quad (6)$$

The resulting objective function is in the standard form of CSC. Thus, efficient algorithms such as coordinate descent (CD) [48], fast iterative soft-thresholding algorithm (FISTA) [57], and feature sign search (FSS) [46] can be used to solve for  $S$ . In our work, we adopt matching pursuit (MP) [58] due to its easy implementation and prominent performance.

### 3.3.2 Optimizing background waveforms of $D_0$

After obtaining the optimal sparse codes  $S$  as described in the previous paragraph, we update the waveforms  $D$  while keeping  $S$  fixed. In order to enforce the separation of state-specific waveforms from the background ones, we propose to learn these two types of waveforms from different subsets of signals. Specifically, we first optimize the background waveforms using signals from all states, with the state-specific waveforms fixed. We then obtain the state-specific waveforms from the signals of the corresponding state, after subtracting the background components.

We first formulate the optimization problem for learning  $D_0$  as follows:

$$\min_{D_0} \left\{ \sum_{c=1}^C \|y_c(t) - g(S_0^c, D_0, t)\|_F + \lambda_c \text{corr}(D_0, D_{-0}) + \lambda_s \text{corr}(D_0, D_0) \right\} \quad \text{s.t.} \quad \|d_k^0\|_2 = 1, \forall k. \quad (7)$$

Here,  $y_c(t)$  denotes the residuals of  $x_c(t)$  after removing the state-specific components:  $y_c(t) = x_c(t) - g(S_c, D_c, t)$ .

For the practical implementation, we randomly update each background waveform  $d_k^0$  from  $D_0$  while keeping resting ones fixed. Then, we define  $y_c^k(t)$  as the component of  $y_c(t)$  that is only composed of  $d_k^0$ :

$$y_c^k(t) = y_c(t) - \sum_{j \in \{1, \dots, w_0\} \setminus k} \sum_{i=1}^{n_j^{0,c}} a_{0,j}^{c,i} d_j^0(t - \tau_{0,j}^{c,i}). \quad (8)$$

To replace the original  $\text{corr}$  with the differentiable  $\tilde{\text{corr}}$ , we first approximate the maximum function smoothly with LogSumExp (LSE) function. Furthermore, we rewrite the discrete convolution between  $d_i$  and  $\tilde{d}_j$  as the matrix multiplication between  $d_i$ 's Toeplitz matrix  $\text{Toe}(d_i)$  and  $\tilde{d}_j$ . Therefore,  $\tilde{\text{corr}}$  can be written as follows:  $\tilde{\text{corr}}(d_i, d_j) = \text{LSE}(\text{Toe}(d_i) \tilde{d}_j)$ . As a result, the problem of solving for a certain background waveform  $d_k^0$  can be formulated as follows:

$$\min_{d_k^0} \left\{ \sum_{c=1}^C \left\| y_c^k(t) - \sum_{i=1}^{n_k^{0,c}} a_{0,k}^{c,i} d_k^0(t - \tau_{0,k}^{c,i}) \right\|_F + \lambda_c \tilde{\text{corr}}(d_k^0, D_{-0}) + \lambda_s \tilde{\text{corr}}(d_k^0, D_0) \right\} \quad \text{s.t.} \quad \|d_k^0\|_2 = 1. \quad (9)$$

To update the background waveform  $d_k^0$ , we utilize stochastic gradient descent (SGD) in our work. Since  $d_k^0$  is sparsely activated and for the sake of computational efficiency, we can optimize only with the patches where  $d_k^0$  is present, following the approach in [31].

### 3.3.3 Optimizing state-specific waveforms of $D_c$

Given fixed sparse codes  $S$  and a background dictionary  $D_0$ , the state-specific waveforms from  $D_c$  (where  $c \in \{1, \dots, C\}$ ) can be optimized by subtracting the background components from  $x_c(t)$ . Let  $z_c(t)$  denote the state-specific component of  $x_c(t)$ :  $z_c(t) = x_c(t) - g(S_0^c, D_0, t)$ . The optimization problem with respect to  $D_c$  is illustrated as follows:

$$\min_{D_c} \{ \|z_c(t) - g(S_c, D_c, t)\|_F + \lambda_c \text{corr}(D_c, D_{-c}) + \lambda_s \text{corr}(D_c, D_c) \} \quad \text{s.t.} \quad \|d_k^c\|_2 = 1, \forall k. \quad (10)$$

The elements of  $D_c$  can be updated randomly using SGD, following the same procedure as described in (9).

### 3.3.4 Implementation details

The overall framework for learning sparse codes and waveforms is shown in Algorithm 1. As discussed above, we optimize sparse codes, background, and state-specific waveforms alternately until the dictionaries converge or the iteration count exceeds a predefined limit. To achieve higher convergence rates, at each iteration, we randomize the order of dictionary elements for the optimization. Moreover, the unit-norm regularizations improve the convergence rate by limiting the search space.

For the sake of computational efficiency, we first divide the band-passed and single-channel input signals into several trials. Each input variable  $x_c$ , where  $c \in \{1, \dots, C\}$ , represents a trial from a certain state. State-specific and background waveforms are then extracted from a set of neural signals containing  $C$

**Algorithm 1** State-sensitive CSC algorithm

---

**Require:** Band-passed and single-channel brain signals of  $C$  states:  $\{x_c(t)\}_{c=1}^C$ ; length of waveforms  $L$ ; weights of sparsity, cross-incoherence, and self-incoherence constraints  $\lambda_w, \lambda_c$ , and  $\lambda_s$ ; size of background and state-specific dictionaries  $w_0$  and  $\{w_c\}_{c=1}^C$ ; time segment length and window overlap ratio of  $\{x_c(t)\}_{c=1}^C$ ; maximum number of iterations  $N_{\text{iter}}$ .

**Ensure:** Background dictionary  $D_0 = \{d_k^0 \in \mathbb{R}^L\}_{k=1}^{w_0}$ ; state-specific dictionaries  $\{D_c = \{d_k^c \in \mathbb{R}^L\}_{k=1}^{w_c}\}_{c=1}^C$ .

- 1: Initialize  $D_0^0$  and  $\{D_c^0\}_{c=1}^C$  with white Gaussian noises, current iteration count it with 0;
- 2: **while** values of  $D_0^{\text{it}}$  and  $\{D_c^{\text{it}}\}_{c=1}^C$  do not converge  $\wedge$  it  $<$   $N_{\text{iter}}$  **do**
- 3:   it  $\leftarrow$  it + 1;
- 4:   {Subsection 3.3.1}
- 5:   **for**  $c = 1$  **to**  $C$  **do**
- 6:     Optimize sparse codes  $\{S_0^c, S_c\}$  via  $x_c(t)$  and  $\{D_0^{\text{it}-1}, D_c^{\text{it}-1}\}$  based on (6) with MP;
- 7:   **end for**
- 8:   {Subsection 3.3.2}
- 9:   **for**  $c = 1$  **to**  $C$  **do**
- 10:     Update  $y_c(t)$  via  $x_c(t)$ ,  $S_c$  and  $D_c^{\text{it}-1}$ ;
- 11:   **end for**
- 12:   **for**  $k$  **in** shuffle( $\{1, \dots, w_0\}$ ) **do**
- 13:     **for**  $c = 1$  **to**  $C$  **do**
- 14:       Update  $y_c^k(t)$  via  $y_c(t)$ ,  $S_0^c$  and  $D_0^{\text{it}-1}$  based on (8);
- 15:     **end for**
- 16:     Optimize  $d_k^0$  via  $\{y_c^k(t)\}_{c=1}^C$  based on (9) with SGD; update the corresponding element of  $D_0^{\text{it}-1}$  via  $d_k^0$ ;
- 17:   **end for**
- 18:    $D_0^{\text{it}} \leftarrow \{d_k^0\}_{k=1}^{w_0}$ ;
- 19:   {Subsection 3.3.3}
- 20:   **for**  $c = 1$  **to**  $C$  **do**
- 21:     Update  $z_c(t)$  via  $x_c(t)$ ,  $S_0^c$  and  $D_0^{\text{it}}$ ;
- 22:     **for**  $k$  **in** shuffle( $\{1, \dots, w_c\}$ ) **do**
- 23:       Update  $z_c^k(t)$  via  $z_c(t)$ ,  $S_c$  and  $D_c^{\text{it}-1}$  in the same way as (8);
- 24:       Optimize  $d_k^c$  via  $z_c^k(t)$  based on (10) with SGD; update the corresponding element of  $D_c^{\text{it}-1}$  via  $d_k^c$ ;
- 25:     **end for**
- 26:      $D_c^{\text{it}} \leftarrow \{d_k^c\}_{k=1}^{w_c}$ ;
- 27:   **end for**
- 28: **end while**
- 29:  $D_0 \leftarrow D_0^{\text{it}}$ ,  $\{D_c \leftarrow D_c^{\text{it}}\}_{c=1}^C$ .

---

trials from  $C$  different states. The resultant state-specific or background waveforms are clustering centres obtained by the modified  $k$ -means algorithm proposed in [59].

Main hyper-parameters of Algorithm 1 include the size of background and state-specific dictionaries  $\{w_c\}_{c=0}^C$ , the weights of the regularization terms on incoherence ( $\lambda_c, \lambda_s$ ), the length of waveforms  $L$ , the time segment length of  $\{x_c(t)\}_{c=1}^C$ , and the window overlap ratio of  $\{x_c(t)\}_{c=1}^C$ . Since MP replaces the sparsity constraint with a fixed constant to control sparsity, the weight  $\lambda_w$  can be omitted in our study of hyper-parameters. In the following experiments, we investigate the effect of these hyper-parameters on the obtained state-specific waveforms. We initialize the waveforms using Gaussian white noises in a standard way. Alternatively, we can use random chunks of the signal after projecting each signal on a rank-1 approximation for waveform initialization [33].

### 3.4 Encoding discriminative features with learned state-specific waveforms

With the resultant waveforms, we encode discriminative features and predict the state of signals to validate their performance as potential biomarkers. The overall framework is illustrated in Figure 2 under the context of seizure prediction.

A feature space consisting of time points ( $\tau$ ) and amplitudes ( $a$ ) can be created using the learned sparse codes [51]. Marked point process (MPP) representation, which allows for the extension of established point process models, has been utilized for the feature extraction [60, 61]. Since state-specific waveforms are assumed to occur frequently only in a certain state, we propose to use the occurrence proportion of the MPP induced by state-specific waveforms as a discriminative feature for detecting state  $c$ . Specifically, the occurrence proportion of  $D_c$  is calculated after decomposing signals with  $D_c$  and  $D_0$ , as described in [61]. Therefore, the discriminative feature is defined as follows:

$$f_c = \frac{\sum_{k=1}^{w_c} \hat{n}_k^c}{\sum_{k=1}^{w_0} \hat{n}_k^0 + \sum_{k=1}^{w_c} \hat{n}_k^c}, \quad (11)$$

where  $\sum_{k=1}^{w_0} \hat{n}_k^0$  and  $\sum_{k=1}^{w_c} \hat{n}_k^c$  represent the total number of phasic events induced by the background dictionary  $D_0$  and state-specific dictionary  $D_c$  in a given trial after the decomposition, respectively.

Based on the discriminative feature  $f_c$ , the state of signals can be predicted. During the training phase, a suitable threshold value for  $f_c$ ,  $\gamma_c$ , is chosen. During the test phase,  $f_c$  is computed using the waveforms learned during training. If the resulting value of  $f_c$  exceeds the predefined threshold  $\gamma_c$ , the trial is deemed to be in state  $c$ , indicating frequent occurrences of state-specific waveforms.

## 4 Experiments and results

### 4.1 Data and settings

#### 4.1.1 Brain signals and preprocessing

We evaluate the performance of our approach in detecting potential prognostic biomarkers with epilepsy brain signals from the Freiburg iEEG dataset [62, 63]. This dataset contains invasive EEG recordings sampled at a frequency of 256 Hz from 21 subjects with intractable epilepsy. Considering that sufficient seizures are required for training, we select 9 subjects solely based on the criterion that the number of seizure onsets is equal or greater than 5 for our experiments. For each patient, recordings from three focal and three extra-focal electrode contacts are available. Signals of the three focal electrode contacts are used for most of our experiments.

Seizure prediction aims to detect the pre-ictal state, which refers to the periods just before the onset of seizures, to prevent epileptic onsets. The Freiburg dataset contains at least 50-min pre-ictal and 24-h interictal recordings per patient on average. In our study, we segment the 15 min before seizure onsets as the pre-ictal state and 60-min discontinuous interictal recordings as the interictal state. Considering that gamma rhythms are highly relevant to epileptic onsets [64], the single-channel raw signals are first filtered into the gamma frequency band (30–100 Hz) using a fifth-order Butterworth filter. After that, they are partitioned into 15-s non-overlapping segments as sets of input trials  $\{x_c\}_{c=1}^C$ .

#### 4.1.2 Train-test split and evaluation metrics

Experiments are carried out on subject-specific seizure prediction via leave-one-out (LOO) validation. Epilepsy patients exhibit notable inter-individual variability in electrode positions stemming from heterogeneity in epileptogenic zones and seizure onset patterns. Consequently, the recorded waveforms between subjects can differ considerably. Therefore, we utilize signals from the same subject for both training and testing to maintain the consistency of electrode positions and recorded waveforms. Specifically, pre-ictal trials related to all but one seizure onset and 80% of interictal trials from a certain subject are employed in the training phase, while the resting trials from the same subject are utilized for the test. Pre-ictal trials are assumed as positive samples, while interictal trials are assumed as negative ones. Since only one onset is present in test trials, the focus of results is on precision, which measures the accuracy of samples when classified as pre-ictal. The precision is computed as follows:  $\text{precision} = \frac{\text{tp}}{\text{tp} + \text{fp}}$ , where tp denotes the number of true positive results, and fp represents the number of false positive results. Note that the precision presented in our results is averaged over 5 random runs of LOO cross-validation, which enhances the reliability and robustness of our findings.

#### 4.1.3 Hyper-parameter settings

The main hyper-parameters of our approach are set as follows. The length of waveforms, denoted as  $L$ , is set to 50 based on the selected gamma band and sampling frequency. The values of the regularization coefficients,  $\lambda_c$  and  $\lambda_s$ , are chosen to be 0.03 and 0.01, respectively. For convenience, we set the size of all state-specific dictionaries,  $\{w_c\}_{c=1}^C$ , to be the same:  $w_c(c \in \{1, \dots, C\}) = 2$ . The size of the background dictionary, represented as  $w_0$ , is set to 5. As described in Subsection 4.1.1, the time segment length and window overlap ratio for  $\{x_c(t)\}_{c=1}^C$  are selected as 15 s and 0%, respectively. The threshold of the classifier,  $\gamma_c$ , for seizure prediction is set to be the 90th percentile of  $f_c$  extracted from training interictal trials. We set the maximum number of iterations  $N_{\text{iter}}$  to 50.

### 4.2 Comparison with existing CSC-based approaches

In order to evaluate the performance of our state-specific waveforms, we compare their performance on subject-specific seizure prediction with that of existing CSC-based approaches. The selected CSC-based approaches are briefly introduced as follows:

**Table 1** Quantitative comparison on prediction precision (%) with existing CSC-based approaches

	pat003			pat004			pat005		
	FE1	FE2	FD1	HR_2	HR_5	TBB1	G_A3	G_A4	IHB4
CSC	63.53 ± 3.46	58.98 ± 1.47	52.59 ± 10.51	77.85 ± 8.25	85.08 ± 3.41	83.52 ± 3.12	38.73 ± 15.48	53.29 ± 8.80	54.38 ± 11.62
CSC-R	63.02 ± 5.37	64.30 ± 4.42	65.10 ± 3.18	<b>86.44</b> ± 2.31	83.86 ± 3.72	87.59 ± 0.77	39.01 ± 13.52	53.98 ± 10.52	45.79 ± 3.85
Ours	<b>76.86</b> ± 2.74	<b>77.49</b> ± 5.56	<b>78.43</b> ± 1.91	84.17 ± 1.97	<b>86.60</b> ± 1.43	<b>90.01</b> ± 2.48	<b>65.54</b> ± 5.71	<b>66.35</b> ± 3.54	<b>67.45</b> ± 5.41
	pat009			pat010			pat017		
	TBA3	TBB2	TBC4	TLA1	TLB1	TLB2	TBA1	TBA2	TLA1
CSC	78.65 ± 2.01	76.43 ± 0.44	74.06 ± 1.13	83.23 ± 3.39	82.80 ± 2.70	70.02 ± 7.17	83.41 ± 8.81	72.73 ± 11.38	91.30 ± 2.36
CSC-R	85.45 ± 4.83	<b>87.64</b> ± 4.29	84.28 ± 4.68	77.43 ± 1.08	<b>83.86</b> ± 3.01	73.62 ± 2.24	96.27 ± 0.33	<b>97.44</b> ± 0.35	<b>96.06</b> ± 0.97
Ours	<b>88.52</b> ± 0.77	87.47 ± 1.80	<b>87.36</b> ± 4.76	<b>87.73</b> ± 1.91	77.29 ± 2.35	<b>85.56</b> ± 1.87	<b>96.51</b> ± 0.51	96.15 ± 0.31	95.75 ± 0.29
	pat018			pat020			pat021		
	G_A7	G_C5	G_E5	G_D3	G_B4	G_A2	TBB1	G_E6	TBA3
CSC	43.88 ± 9.67	33.25 ± 11.64	38.10 ± 13.02	66.56 ± 4.87	57.26 ± 4.64	45.08 ± 7.53	45.27 ± 11.16	65.16 ± 6.13	70.74 ± 5.59
CSC-R	35.82 ± 7.63	43.83 ± 10.29	55.92 ± 3.99	61.72 ± 6.89	58.10 ± 4.41	47.38 ± 8.78	61.60 ± 6.95	73.87 ± 9.07	<b>76.64</b> ± 2.60
Ours	<b>88.00</b> ± 3.06	<b>91.23</b> ± 1.91	<b>86.15</b> ± 5.77	<b>73.05</b> ± 5.23	<b>65.22</b> ± 4.04	<b>79.17</b> ± 3.49	<b>84.76</b> ± 0.51	<b>74.48</b> ± 3.91	69.99 ± 4.55

- **CSC.** The standard CSC learns recurrent shift-invariant waveforms in an unsupervised manner. We use the MP-SVD algorithm [31] to extract prototypical waveforms from pre-ictal and interictal trials. We designate the waveforms obtained from pre-ictal trials as the pre-ictal-specific dictionary  $D_{\text{pre-ictal}}$ , while the waveforms obtained from interictal trials form the background dictionary  $D_0$ .

- **CSC-R.** One criterion for identifying “task-related” waveforms is an increase in the activation rate during task-performance [36]. We incorporate this criterion into CSC-R for selecting state-specific waveforms. CSC-R involves learning waveforms from pre-ictal trials using standard CSC, and then selecting waveforms with a higher rate of increase in average  $f_{\text{pre-ictal}}$  from interictal to pre-ictal periods post-hoc. These selected waveforms form the pre-ictal-specific dictionary, while the background dictionary consists of resting waveforms learned from pre-ictal trials and all waveforms extracted from interictal trials.

We choose CSC and CSC-R as the competitor based on the following considerations. First, CSC is a data-driven method that has demonstrated better performance than the knowledge-based ones due to its lesser dependence on predefined assumptions. Among these data-driven methods, CSC has emerged as an efficient tool due to its flexibility in incorporating physical priors [36]. Second, recent CSC-based studies have focused on “task-related” waveforms, similar to our goal of learning state-specific biomarkers. As mentioned in Subsection 2.2, these studies conduct additional selections after the standard CSC, of which CSC-R is a representative one suitable for biomarker detections.

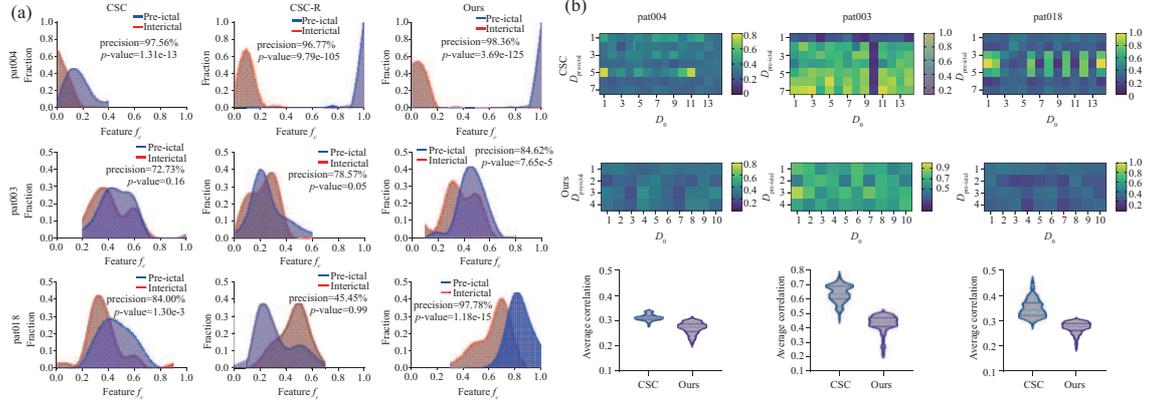
#### 4.2.1 Quantitative results

The average prediction precision of three focal electrode contacts from 9 subjects is presented in Table 1. The results indicate that our method achieves the highest precision in most cases. Notably, our method outperforms the baselines by over 10.00% precision on subjects such as pat005, pat018, and pat020, highlighting the importance of incorporating state-specificity directly into the learning process for identifying potential biomarkers over learning in an unsupervised manner and with the additional selection. Additionally, our method achieves precision no worse than 5.00% on average on channels where CSC-R performs the best. We also observe that CSC-R achieves better precision in most cases compared with standard CSC, indicating the effectiveness of additional selections based on  $f_c$ .

#### 4.2.2 Visualization results

To further explore the reasons for our superiority in potential biomarker identification, we conduct visualizations of  $f_c$  and analyze the correlation between state-specific and background waveforms. Figure 3 presents the results of this analysis for three selected subjects (pat004, pat003, and pat018).

Subsection 3.4 mentions that test epileptic signals are classified based on the activation rate of state-specific waveforms from  $D_{\text{pre-ictal}}$ . Therefore, pre-ictal signals are supposed to be higher in  $f_c$ , which suggests the higher activation rate of pre-ictal-specific waveforms. Figure 3(a) demonstrates a comparison of the frequency distribution of  $f_c$  using waveforms extracted by CSC, CSC-R, and our method. Our method and CSC-R exhibit a more significant differentiation in the distribution of  $f_c$  than standard CSC for pat004. When comparing our performance with CSC and CSC-R on the other two selected subjects,



**Figure 3** (Color online) Results of the visualization analysis performed on three selected subjects (pat004, pat003, and pat018). (a) Distribution of  $f_c$  computed on test pre-ictal and interictal trials using waveforms extracted by three different methods: CSC, CSC-R, and our method. Each figure includes the corresponding prediction precision and the p-value of one-tailed t-test conducted on  $f_c$  between test pre-ictal and interictal trials. (b) Correlation matrix visualization and violin plots of the average correlation between output  $D_{pre-ictal}$  and  $D_0$  learned by CSC and our method during one training. The correlation value displayed in the violin plot represents the average correlation between output  $D_{pre-ictal}$  and  $D_0$  learned from a set of input trials.

we find that our distribution peaks are the most distant, giving rise to the highest prediction precision. It is worth noting that CSC-R demonstrates reverse peaks in the distribution of pat018, indicating the instability of the additional selection based on average increase rates. These observations suggest that the background waveforms can obscure the distribution differences of  $f_c$  between pre-ictal and interictal, and our method improves the prediction precision by boosting these differences.

The correlations between the learned  $D_{pre-ictal}$  and  $D_0$  are computed using the function `corr`, and the further investigation on the correlations is carried out to understand the superior performance of our state-specific waveforms. Figure 3(b) depicts part of the correlation matrix and the distribution of the average correlation between state-specific and background waveforms. Bright patches that signify great similarities are present in all correlation matrixes of CSC. In contrast, no obvious bright patches are visible in the visualization of our correlation matrixes. Consequently, these results confirm the existence of background waveforms that appear randomly throughout all states. Additionally, the visualization suggests that our method captures the background waveforms, which indicates the successful separation of state-specific prototypical waveforms from background ones. Similarly, the violin plots displayed in the last column of Figure 3(b) show that the average correlation of CSC is generally higher than ours. This trend demonstrates the existence of background waveforms and highlights the advantage of identifying potential prognostic biomarkers via state-specific waveforms.

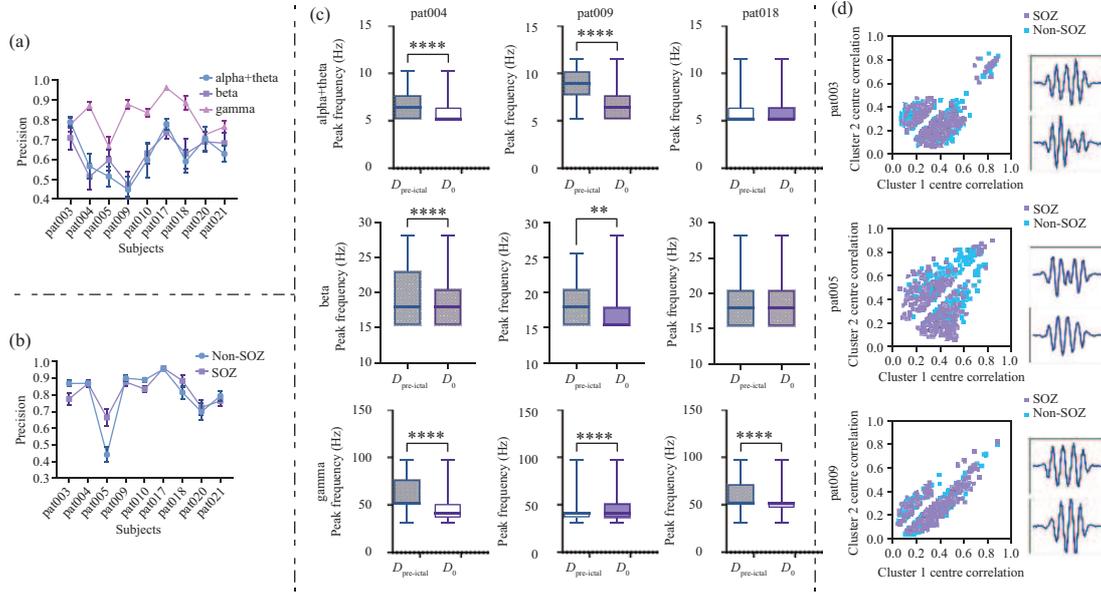
### 4.3 Post analysis on learned state-specific waveforms

Besides examining the performance of pre-ictal-specific waveforms in seizure prediction, post analysis is conducted to interpret these state-specific waveforms. Specifically, we perform experiments on state-specific waveforms across different frequency bands and channels to better understand the characteristics of our learned potential biomarkers.

#### 4.3.1 Exploring learned waveforms in different frequency bands

The raw iEEG recordings are typically divided into several classical frequency bands, namely theta (4–8 Hz), alpha (8–12 Hz), beta (13–30 Hz), and gamma (30–100 Hz). In our experiment, we merge the theta and alpha bands to form a new frequency band, which we refer to as “alpha+theta”. The analytical results of our learned waveforms from various frequency bands are summarized in Figures 4(a) and (c), and Figure 5.

The precision of subject-specific prediction across different frequency bands is presented in Figure 4(a). The quantitative results reveal that state-specific waveforms extracted from the gamma frequency band outperform those from the other two frequency bands for most subjects. This is consistent with the findings of previous studies [65], which have demonstrated that increased gamma connections in epileptic tissue can differentiate it from healthy tissue.



**Figure 4** (Color online) (a) Subject-specific prediction precision using state-specific waveforms extracted from different frequency bands. (b) Average subject-specific prediction precision of SOZ and non-SOZ across nine chosen subjects. (c) Boxplots showing the peak frequency distributions of waveforms from  $D_{\text{pre-ictal}}$  and  $D_0$  learned by our method from different frequency bands. Results marked with “\*\*\*” and “\*\*\*\*” indicate that the  $p$ -values obtained from the unpaired one-tailed  $t$ -test are below  $1e-2$  and  $1e-4$ , respectively. (d) The two-dimensional correlation representation of state-specific waveforms identified in SOZ and non-SOZ in one training. The waveforms on the right-hand side of the panel are the visualization of each cluster centre.

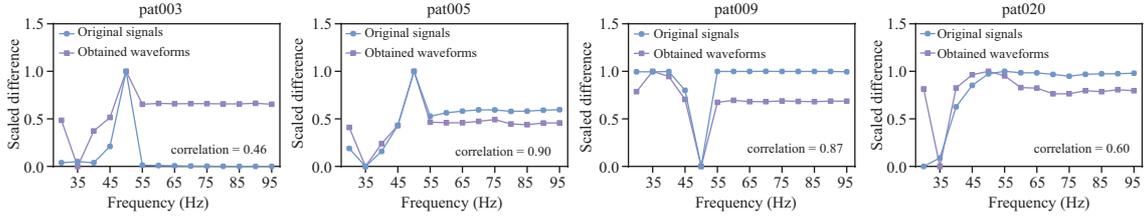
To conduct a thorough examination of the changes in characteristics of waveforms from interictal to pre-ictal periods, we compare the peak frequency distribution of state-specific waveforms from  $D_{\text{pre-ictal}}$  with that of background waveforms from  $D_0$  during a certain training. Figure 4(c) displays the boxplots for three selected subjects (pat004, pat009, and pat018). Each plot consists of boxes that span between the first and third quartiles (Q1 and Q3), lines at the median, and whiskers that extend to the highest and lowest data points. The differences in the distributions are statistically significant for all three subjects in the gamma band. In contrast, for pat018, the distributions of  $D_{\text{pre-ictal}}$  are similar to those of  $D_0$  in the alpha+theta and beta frequency bands. Furthermore, trends of constant, increasing, and decreasing peak frequencies from state-specific to background waveforms are observed, making it challenging to differentiate potential prognostic biomarkers solely based on the peak frequency. Overall, these observations suggest that changes in peak frequency are most pronounced in the gamma frequency band and are in line with previous findings that gamma rhythms are related to epileptic seizure onsets [64].

To further explore the correlation in high-frequency components between the obtained dictionaries and original signals, we first calculate the difference in the peak frequency distribution between waveforms from the pre-ictal-specific dictionary  $D_{\text{pre-ictal}}$  and those from the shared dictionary  $D_0$ . We also calculate the difference in power spectral densities between pre-ictal and interictal trials using Welch’s method. Both differences are computed within the gamma frequency band and scaled to a range between 0 and 1. Figure 5 shows the scaled differences in peak frequency distributions and power spectral densities for selected subjects pat003, pat005, pat009, and pat020. The former difference correlates strongly with the latter, suggesting that the distinguishing features of the state-specific waveforms arise from shifts in the high-frequency content of the original signals across different states.

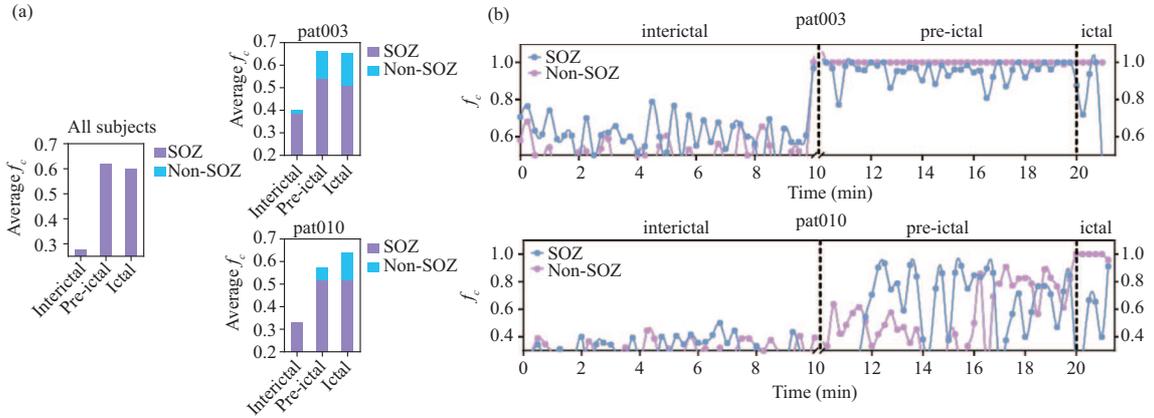
#### 4.3.2 Exploring learned waveforms in different channels

The seizure onset zone (SOZ) [66,67] refers to as the region in the brain where epileptic seizures originate. In this experiment, we aim to figure out whether the position of the channel affects the characteristics of detected potential prognostic biomarkers. The primary outcomes of the study are summarized in Figures 4(b) and (d), and Figure 6.

Initially, we compare the average subject-specific prediction precision of SOZ with that of non-SOZ. As demonstrated in Figure 4(b), the state-specific waveforms obtained from SOZ exhibit similar prediction precision to those derived from non-SOZ in most subjects. This observation indicates that state-specific



**Figure 5** (Color online) Difference in peak frequency distributions between waveforms from the pre-ictal-specific dictionary  $D_{\text{pre-ictal}}$  and the shared dictionary  $D_0$ . It also shows the difference in power spectral densities between pre-ictal and interictal trials. Both differences are computed within the gamma frequency band and scaled to a range from 0 to 1. Finally, the figure depicts the Pearson correlation coefficient between these two scaled differences for subjects pat003, pat005, pat009, and pat020.

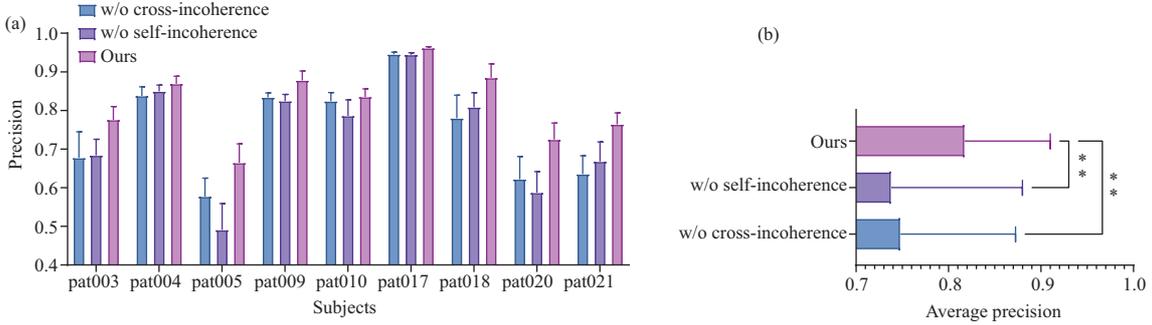


**Figure 6** (Color online) Portion of results from post analysis conducted on state-specific waveforms obtained from SOZ and non-SOZ. (a) Comparison on the average  $f_c$  values between interictal, pre-ictal, and ictal trials of SOZ and non-SOZ for all subjects and two selected subjects. (b) The detailed changes in  $f_c$  from interictal to pre-ictal, and ictal trials in a single channel located inside and outside the SOZ, respectively, for pat003 and pat010. The interictal and pre-ictal trials are disjointed in the figure, whereas the pre-ictal and ictal trials are contiguous.

waveforms exist in both SOZ and non-SOZ.

To evaluate the similarity between state-specific waveform patterns from SOZ and non-SOZ, we employ dimensionality reduction techniques to directly represent these waveforms from the same subject. Specifically, based on the default value of 2 for  $w_c$ , we cluster the state-specific waveforms across channels into two clusters via the modified  $k$ -means algorithm [59]. We then compute the correlation of each waveform with resultant cluster centers using the function `corr` and obtain the two-dimensional representation. The finding, as shown in Figure 4(d) for three representative subjects (pat003, pat005, and pat009), demonstrates that the state-specific waveforms present in SOZ and non-SOZ exhibit similar patterns. This observation indicates that these similar waveforms coexist in both regions. The diversity of state-specific waveforms shown on the right side of the figure indicates the subject-specificity of seizure prediction, a finding consistent with prior researches [68, 69].

In order to ascertain the co-existence of potential biomarkers in both SOZ and non-SOZ, more experiments are conducted to examine the changes in their activation proportions during interictal, pre-ictal, and ictal trials. The results displayed in Figure 6(a) indicate that the average  $f_c$  of the SOZ and non-SOZ exhibits similar trends. Across all subjects, there is a substantial increase in the average  $f_c$  from interictal to pre-ictal trials, indicating that the detected waveforms are specific to the pre-ictal periods. Interestingly, the average  $f_c$  remains high in ictal trials despite the lack of such trials in the training phase, suggesting the frequent occurrence of pre-ictal-specific waveforms in ictal trials as well. The detailed changes of two selected subjects (pat003 and pat010) are illustrated in Figure 6(b), which reveal sustained low  $f_c$  in interictal signals, followed by a sharp increase in  $f_c$  during pre-ictal periods, and sustained high  $f_c$  in ictal signals. These observations are consistent with the trends observed in the percentages of time occupied by ripples, as depicted in Figure 3 of [13].



**Figure 7** (Color online) (a) Comparison of prediction precision between our method and a formulation without the cross-incoherence or self-incoherence regularization term for each subject. (b) Comparison of average precision performance between our method and a formulation without the cross-incoherence or self-incoherence regularization term. The results marked with “\*\*\*” indicate significant differences based on paired one-way analysis of variance (ANOVA).

#### 4.4 Effect of hyper-parameter settings

As delineated in Subsection 3.3.4, our approach uses a set of hyper-parameters. These hyper-parameters include the weights of incoherence regularization terms ( $\lambda_s$ ,  $\lambda_c$ ), the size of the background dictionary  $w_0$ , the size of the state-specific dictionary  $\{w_c\}_{c=1}^C$ , the length of waveforms which is measured in units of time points and denoted by  $L$ , the time segment length and the window overlap ratio of  $\{x_c(t)\}_{c=1}^C$ . To investigate the effect of hyper-parameters to the performance of our approach, we conduct a thorough experiment on each parameter based on the prediction precisions of nine selected subjects, while fixing the other hyper-parameters.

##### 4.4.1 Effect of incoherence regularization terms

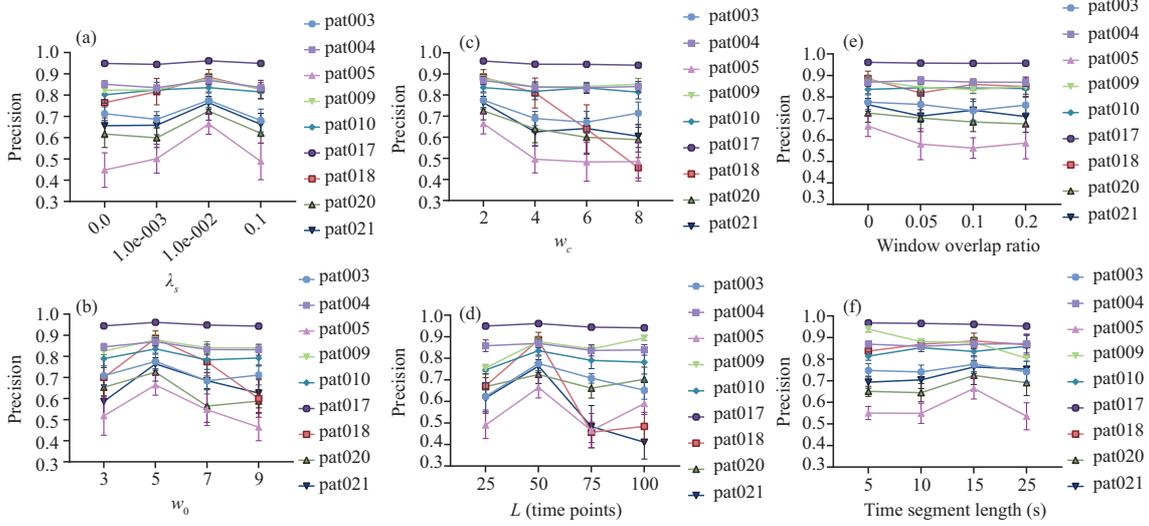
An ablation experiment is conducted to verify the effectiveness of regularization terms. Our method is compared with a formulation without the cross-incoherence or self-incoherence regularization term. As shown in Figure 7(a), our method outperforms the other two methods in terms of prediction precision for all chosen subjects. On average, our method performs at least 5.00% better than both methods, and this result is found to be statistically significant by one-way ANOVA as illustrated in Figure 7(b). These findings demonstrate the usefulness of our proposed cross-incoherence and self-incoherence regularization terms in improving the prediction precision.

We further investigate the impact of the regularization terms by varying their weights. In this study, we set the value of  $\lambda_c$  to three times that of  $\lambda_s$ , namely  $\lambda_c = 3\lambda_s$ , to focus more on the incoherence between different dictionaries. The remaining hyper-parameters are set to the following constants:  $w_0 = 5$ ,  $\{w_c\}_{c=1}^2 = 2$ , and  $L = 50$ . The time segment length and the window overlap ratio are set to 15 s and 0%, respectively. The value of  $\lambda_s$  is varied from 0 to 0.1 and the resulting precision values are presented in Figure 8(a). Our analysis reveals that the optimal value for  $\lambda_s$  is 0.01, which highlights the efficacy of incoherence regularization terms in improving the performance of our learned waveforms.

##### 4.4.2 Effect of dictionary sizes and waveform lengths

The size of the dictionary is a crucial hyper-parameter in convolutional dictionary learning, as it directly impacts the quality of the learned representations. In this experiment, we conduct evaluations on the sizes for background and state-specific dictionaries, denoted by  $w_0$  and  $\{w_c\}_{c=1}^2$ , respectively. We begin by analyzing the optimal size of the background dictionary  $w_0$ , with the remaining hyper-parameters set to the constants mentioned above. As shown in Figure 8(b), the background dictionary with 5 waveforms demonstrates the best precision performance. We also find that a background dictionary with insufficient or excessive waveforms leads to loss or redundancy in representation. A similar result is obtained for the size of state-specific dictionaries, which are set to the same size for convenience. From Figure 8(c), we find the optimal size of the state-specific dictionaries to be 2.

Considering a sampling frequency of 256 Hz and the frequency range of interest being the gamma frequency band (30–100 Hz), the length of waveforms is constrained to fall within the range of 25 to 100 time points. Based on the results shown in Figure 8(d), it is found that 50-time-point-long waveforms, which correspond to about 200 ms, offer the richest information for epilepsy.



**Figure 8** (Color online) Effect of several hyper-parameters on the prediction precision of our approach. (a) The weight of the regularization term  $\lambda_s$ ; (b) the size of the background dictionary  $w_0$ ; (c) the size of the state-specific dictionaries  $w_c$ ; (d) the length of waveforms  $L$ ; (e) the window overlap ratio; (f) the time segment length.

#### 4.4.3 Effect of window overlap ratios and time segment lengths

Experiments are conducted by varying the window overlap ratio from 0% to 20%, as well as the time segment length from 5 to 25 s. To ensure a fair comparison, we simultaneously change the time segment lengths for training and testing. As illustrated in Figure 8(e), the prediction precision is robust over different selections of window overlap ratios for most subjects. For subjects pat005 and pat020, a higher overlap ratio tends to decrease the precision. This could be because noise repeats across adjacent windows, weakening the impact of state-specific waveforms. As shown in Figure 8(f), most subjects achieve similar or better performance as time segment length increases. A possible explanation for this is that longer time segments may capture more stable state-specific patterns.

## 4.5 Discussion

While experimental results demonstrate the superior performance of our state-specific waveforms in seizure prediction, it is important to acknowledge the limitations of our proposed method.

Firstly, alternative feature extraction and optimization methods can be explored. In order to detect a certain state, we extract features based on the occurrence proportion of state-specific waveforms. Although we present one method for encoding signals with state-specific waveforms, there may be other methods for feature extraction in point processes that could be applied to our proposed framework. Thus, further investigations are warranted to explore alternative approaches to feature extraction that can fully capture the latent complexity of the neural signals and enhance the representation of neural signals to improve the accuracy of biomarker identification. Similarly, alternative optimization methods could also be suitable. For instance, we could implement the iterative optimization process with an autoencoder based on neural networks, similar to the fast iterative shrinkage thresholding algorithm (FISTA) [70]. Such autoencoders apply the shrinkage to gradient-based updates at each iteration, and the waveforms can be obtained with their convolutional kernels.

Secondly, due to space constraints, we only validate our approach for potential epileptic biomarkers. However, it should be noted that our proposed approach has the potential to be extended to learn state-specific waveforms as potential prognostic biomarkers for other neurological disorders. For instance, distinctive sharp beta waveforms are observed in invasive motor cortex recordings from Parkinson's disease patients [71, 72]. Likewise, disrupted sleep spindles with decreased amplitudes and durations have the potential to serve as an effective biomarker for the Alzheimer's disease [73]. Moreover, the increase in waveform complexities observed when the brain transitions from anesthesia to wakefulness may assist in diagnosing disorders of consciousness [74, 75]. Due to the state-sensitivity of these waveforms, our approach holds great potential for detecting them as biomarkers. From a clinical perspective, monitoring changes in these waveforms across different states offers a more sensitive method for disease tracking [76].

Additionally, the identification of state-specific waveform alterations can provide valuable insights for personalized treatment strategies. Finally, comparing waveforms before and after therapy across different states allows for a comprehensive assessment of treatment efficacy and can help guide treatment adjustments if needed. Therefore, further studies can focus on fully elucidating the potential clinical applications of these learned state-specific waveforms.

Finally, it should be noted that our proposed framework only utilizes single-channel input trials, which implies that the spatial information of channels is not considered in the current implementation. This may lead to a loss of valuable information, especially in cases where the prognostic biomarkers are distributed across different brain regions. In order to fully capture the spatial characteristics of the prognostic biomarkers, it is necessary to investigate the latent relationship of MPP representations across channels. One potential solution is to explore multivariate point process models that incorporate spatial dependency across different channels [77, 78]. Such models could enable us to extract more informative features and identify potential biomarkers that are specific to certain brain regions or networks. Therefore, future studies should focus on incorporating spatial information into the framework and exploring the potential of multivariate models in improving the accuracy and reliability of potential prognostic biomarkers.

## 5 Conclusion

In this study, we propose a novel CSC-based approach to learn state-specific waveforms from brain signals as candidates for prognostic biomarkers. The core of this approach is to incorporate the state-specificity explicitly into the modeling and learning process. Based on the model formulation with state-specific and background waveforms, a novel state-sensitive CSC algorithm is introduced to learn these waveforms. Quantitative and analytical results on epilepsy brain signals demonstrate that our approach can effectively identify state-specific prototypical waveforms in pre-ictal states, which can potentially be biomarkers for seizure prediction. Overall, our study provides a novel automatic tool for identifying potential prognostic biomarkers via learning state-specific waveforms, which has the potential to be extended to diverse diagnostic predictions and clinical decision-making.

**Acknowledgements** This work was supported by STI 2030 Major Projects (Grant No. 2021ZD0200400), National Natural Science Foundation of China (Grant Nos. U1909202, 61925603, 62276228), and Key R&D Program of Zhejiang (Grant No. 2023C03001).

## References

- 1 Marshall L, Helgadóttir H, Mölle M, et al. Boosting slow oscillations during sleep potentiates memory. *Nature*, 2006, 444: 610–613
- 2 Lundqvist M, Rose J, Herman P, et al. Gamma and beta bursts underlie working memory. *Neuron*, 2016, 90: 152–164
- 3 Wessel J R.  $\beta$ -bursts reveal the trial-to-trial dynamics of movement initiation and cancellation. *J Neurosci*, 2020, 40: 411–423
- 4 Ferrarelli F, Huber R, Peterson M J, et al. Reduced sleep spindle activity in schizophrenia patients. *Am J Psychiatry*, 2007, 164: 483–492
- 5 Wamsley E J, Tucker M A, Shinn A K, et al. Reduced sleep spindles and spindle coherence in schizophrenia: mechanisms of impaired memory consolidation? *Biol Psychiatry*, 2012, 71: 154–161
- 6 Limoges É, Mottron L, Bolduc C, et al. Atypical sleep architecture and the autism phenotype. *Brain*, 2005, 128: 1049–1061
- 7 Sun X, Qi Y, Wang Y, et al. Convolutional multiple instance learning for sleep spindle detection with label refinement. *IEEE Trans Cogn Dev Syst*, 2022, 15: 272–284
- 8 Shibagaki M, Kiyono S, Watanabe K. Spindle evolution in normal and mentally retarded children: a review. *Sleep*, 1982, 5: 47–57
- 9 Li S, Zhou W, Yuan Q, et al. Seizure prediction using spike rate of intracranial EEG. *IEEE Trans Neural Syst Rehabil Eng*, 2013, 21: 880–886
- 10 Fukumori K, Yoshida N, Sugano H, et al. Epileptic spike detection using neural networks with linear-phase convolutions. *IEEE J Biomed Health Inform*, 2021, 26: 1045–1056
- 11 Chahid A, Albalawi F, Alotaiby T N, et al. QuPWM: feature extraction method for epileptic spike classification. *IEEE J Biomed Health Inform*, 2020, 24: 2814–2824
- 12 Jacobs J, Zijlmans M. HFO to measure seizure propensity and improve prognostication in patients with epilepsy. *Epilepsy Curr*, 2020, 20: 338–347
- 13 Zijlmans M, Jacobs J, Kahn Y U, et al. Ictal and interictal high frequency oscillations in patients with focal epilepsy. *Clin NeuroPhysiol*, 2011, 122: 664–671
- 14 Bai D, Liu T, Han X, et al. Application research on optimization algorithm of sEMG gesture recognition based on light CNN+LSTM model. *Cyborg Bionic Syst*, 2021, 2021:
- 15 Shin H, Law R, Tsutsui S, et al. The rate of transient beta frequency events predicts behavior across tasks and species. *eLife*, 2017, 6: e29086
- 16 Brady B, Power L, Bardouille T. Age-related trends in neuromagnetic transient beta burst characteristics during a sensorimotor task and rest in the Cam-CAN open-access dataset. *NeuroImage*, 2020, 222: 117245
- 17 Caplan J B, Bottomley M, Kang P, et al. Distinguishing rhythmic from non-rhythmic brain activity during rest in healthy neurocognitive aging. *NeuroImage*, 2015, 112: 341–352

- 18 Kosciessa J Q, Grandy T H, Garrett D D, et al. Single-trial characterization of neural rhythms: potential and challenges. *NeuroImage*, 2020, 206: 116331
- 19 Fabus M S, Quinn A J, Warnaby C E, et al. Automatic decomposition of electrophysiological data into distinct nonsinusoidal oscillatory modes. *J NeuroPhysiol*, 2021, 126: 1670–1684
- 20 Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc R Soc Lond A*, 1998, 454: 903–995
- 21 Cole S, Voytek B. Cycle-by-cycle analysis of neural oscillations. *J NeuroPhysiol*, 2019, 122: 849–861
- 22 Abeles M. Revealing instances of coordination among multiple cortical areas. *Biol Cybern*, 2014, 108: 665–675
- 23 Tal I, Abeles M. Temporal accuracy of human cortico-cortical interactions. *J NeuroPhysiol*, 2016, 115: 1810–1820
- 24 Tal I, Abeles M. Imaging the spatiotemporal dynamics of cognitive processes at high temporal resolution. *Neural Comput*, 2018, 30: 610–630
- 25 Hyvärinen A, Oja E. Independent component analysis: algorithms and applications. *Neural Netw*, 2000, 13: 411–430
- 26 Vidaurre D, Quinn A J, Baker A P, et al. Spectrally resolved fast transient brain states in electrophysiological data. *NeuroImage*, 2016, 126: 81–95
- 27 Becker R, Vidaurre D, Quinn A J, et al. Transient spectral events in resting state MEG predict individual task responses. *NeuroImage*, 2020, 215: 116818
- 28 Coquelet N, de Tiège X, Roshchupkina L, et al. Microstates and power envelope hidden Markov modeling probe bursting brain activity at different timescales. *NeuroImage*, 2022, 247: 118850
- 29 Michel C M, Koenig T. EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. *NeuroImage*, 2018, 180: 577–593
- 30 Jost P, Vandergheynst P, Lesage S, et al. Motif: an efficient algorithm for learning translation invariant dictionaries. In: *Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing*, Toulouse, 2006. 857–860
- 31 Brockmeier A J, Principe J C. Learning recurrent waveforms within EEGs. *IEEE Trans Biomed Eng*, 2016, 63: 43–54
- 32 Jas M, Tour T D, Simsekli U, et al. Learning the morphology of brain signals using alpha-stable convolutional sparse coding. In: *Proceedings of Advances in Neural Information Processing Systems*, Long Beach, 2017. 1099–1108
- 33 Tour T D, Moreau N, Jas M, et al. Multivariate convolutional sparse coding for electromagnetic brain signals. In: *Proceedings of Advances in Neural Information Processing Systems*, Montréal, 2018. 1–11
- 34 Sherman M A, Lee S, Law R, et al. Neural mechanisms of transient neocortical beta rhythms: converging evidence from humans, computational modeling, monkeys, and mice. *Proc Natl Acad Sci USA*, 2016, 113: 4885–4894
- 35 Freeman W, Quiroga R Q. *Imaging Brain Function with EEG: Advanced Temporal and Spatial Analysis of Electroencephalographic Signals*. New York: Springer Science & Business Media, 2012
- 36 Power L, Allain C, Moreau T, et al. Using convolutional dictionary learning to detect task-related neuromagnetic transients and ageing trends in a large open-access dataset. *NeuroImage*, 2023, 267: 119809
- 37 Allain C, Gramfort A, Moreau T. Dripp: driven point processes to model stimuli induced patterns in M/EEG signals. In: *Proceedings of the 10th International Conference on Learning Representations*, 2022. 1–25
- 38 Jiao Y, Zhang Y, Chen X, et al. Sparse group representation model for motor imagery EEG classification. *IEEE J Biomed Health Inform*, 2018, 23: 631–641
- 39 Huang J, Zhu Q, Hao X, et al. Identifying resting-state multifrequency biomarkers via tree-guided group sparse learning for schizophrenia classification. *IEEE J Biomed Health Inform*, 2018, 23: 342–350
- 40 Zhang Q, Li B X. Discriminative k-SVD for dictionary learning in face recognition. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, 2010. 2691–2698
- 41 Jiang Z L, Lin Z, Davis L S. Learning a discriminative dictionary for sparse coding via label consistent k-SVD. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2011. 1697–1704
- 42 Cheng G, Lai P, Gao D, et al. Class attention network for image recognition. *Sci China Inf Sci*, 2023, 66: 132105
- 43 Iqbal A, Seghouane A K, Adali T. Shared and subject-specific dictionary learning (ShSSDL) algorithm for multisubject fMRI data analysis. *IEEE Trans Biomed Eng*, 2018, 65: 2519–2528
- 44 Gao S H, Tsang I W H, Ma Y. Learning category-specific dictionary and shared dictionary for fine-grained image categorization. *IEEE Trans Image Process*, 2014, 23: 623–634
- 45 Zhao S, Han J, Lv J, et al. Supervised dictionary learning for inferring concurrent brain networks. *IEEE Trans Med Imag*, 2015, 34: 2036–2045
- 46 Grosse R, Raina R, Kwong H, et al. Shift-invariant sparse coding for audio classification. In: *Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence*, Vancouver, 2007. 149–158
- 47 Pachitariu M, Packer A M, Pettit N, et al. Extracting regions of interest from biological images with convolutional sparse block coding. In: *Proceedings of Advances in Neural Information Processing Systems*, 2013. 1–9
- 48 Kavukcuoglu K, Sermanet P, Boureau Y L, et al. Learning convolutional feature hierarchies for visual recognition. In: *Proceedings of Advances in Neural Information Processing Systems*, Vancouver, 2010. 1–9
- 49 Zeiler M D, Krishnan D, Taylor G W, et al. Deconvolutional networks. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, 2010. 2528–2535
- 50 Brockmeier A J. *Learning and exploiting recurrent patterns in neural data*. Dissertation for Ph.D. Degree. Gainesville: University of Florida, 2014
- 51 Akella S, Mohebi A, Principe J C, et al. Marked point process representation of oscillatory dynamics underlying working memory. *J Neural Eng*, 2021, 18: 026016
- 52 Buzsáki G. *Rhythms of the Brain*. New York: Oxford University Press, 2006
- 53 Loza C A, Reddy C G, Akella S, et al. Discrimination of movement-related cortical potentials exploiting unsupervised learned representations from ECoGs. *Front Neurosci*, 2019, 13: 1248
- 54 Ramirez I, Sprechmann P, Sapiro G. Classification and clustering via dictionary learning with structured incoherence and shared features. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, 2010. 3501–3508
- 55 Kong S, Wang D. A dictionary learning approach for classification: separating the particularity and the commonality. In: *Proceedings of European Conference on Computer Vision*, Heraklion, 2012. 186–199
- 56 Ramirez I, Lecumberry F, Sapiro G. Universal priors for sparse modeling. In: *Proceedings of IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, Aruba, 2009. 197–200
- 57 Chalasani R, Principe J C, Ramakrishnan N. A fast proximal method for convolutional sparse coding. In: *Proceedings of International Joint Conference on Neural Networks*, Dallas, 2013. 1–5

- 58 Mallat S G, Zhang Z F. Matching pursuits with time-frequency dictionaries. *IEEE Trans Signal Process*, 1993, 41: 3397–3415
- 59 Charalampidis D. A modified k-means algorithm for circular invariant clustering. *IEEE Trans Pattern Anal Machine Intell*, 2005, 27: 1856–1865
- 60 Huang Y, Zhang X, Shen X, et al. Extracting synchronized neuronal activity from local field potentials based on a marked point process framework. *J Neural Eng*, 2022, 19: 046043
- 61 Loza C A, Okun M S, Príncipe J C. A marked point process framework for extracellular electrical potentials. *Front Syst Neurosci*, 2017, 11: 95
- 62 Winterhalder M, Maiwald T, Voss H U, et al. The seizure prediction characteristic: a general framework to assess and compare seizure prediction methods. *Epilepsy Behav*, 2003, 4: 318–325
- 63 Aschenbrenner-Scheibe R. How well can epileptic seizures be predicted? An evaluation of a nonlinear method. *Brain*, 2003, 126: 2616–2626
- 64 Hughes J R. Gamma, fast, and ultrafast waves of the brain: their relationships with epilepsy and behavior. *Epilepsy Behav*, 2008, 13: 25–31
- 65 Zweiphenning W J E M, Keijzer H M, van Diessen E, et al. Increased gamma and decreased fast ripple connections of epileptic tissue: a high-frequency directed network approach. *Epilepsia*, 2019, 60: 1908–1920
- 66 Kahane P, Minotti L, Hoffmann D, et al. Invasive EEG in the definition of the seizure onset zone: depth electrodes. *Handbook Clin Neurophys*, 2003, 3: 109–133
- 67 Burns S P, Santaniello S, Yaffe R B, et al. Network dynamics of the brain and influence of the epileptic seizure onset zone. *Proc Natl Acad Sci USA*, 2014, 111: 5321–5330
- 68 Karoly P J, Freestone D R, Boston R, et al. Interictal spikes and epileptic seizures: their relationship and underlying rhythmicity. *Brain*, 2016, 139: 1066–1078
- 69 Dissanayake T, Fernando T, Denman S, et al. Geometric deep learning for subject independent epileptic seizure prediction using scalp EEG signals. *IEEE J Biomed Health Inform*, 2021, 26: 527–538
- 70 Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J Imag Sci*, 2009, 2: 183–202
- 71 Cole S R, van der Meij R, Peterson E J, et al. Nonsinusoidal beta oscillations reflect cortical pathophysiology in Parkinson's disease. *J Neurosci*, 2017, 37: 4830–4840
- 72 Jackson N, Cole S R, Voytek B, et al. Characteristics of waveform shape in Parkinson's disease detected with scalp electroencephalography. *eNeuro*, 2019, 6: 1–11
- 73 Weng Y Y, Lei X, Yu J. Sleep spindle abnormalities related to Alzheimer's disease: a systematic mini-review. *Sleep Med*, 2020, 75: 37–44
- 74 Liang Y, Liang J, Song C, et al. Complexity of cortical wave patterns of the wake mouse cortex. *Nat Commun*, 2023, 14: 1434
- 75 Sun X, Qi Y, Ma X, et al. Conformer: consciousness detection using transformer networks with correntropy-based measures. *IEEE Trans Neural Syst Rehabil Eng*, 2023, 31: 2933–2943
- 76 Babaei N, Hannani N, Dabanloo N J, et al. A systematic review of the use of commercial wearable activity trackers for monitoring recovery in individuals undergoing total hip replacement surgery. *Cyborg Bionic Syst*, 2022, 2022: 1–16
- 77 Li W H, Qi Y, Pan G. Online neural sequence detection with hierarchical dirichlet point process. In: *Proceedings of Advances in Neural Information Processing Systems*, 2022. 6654–6665
- 78 Qian C, Sun X, Wang Y, et al. Binless kernel machine: modeling spike train transformation for cognitive neural prostheses. *Neural Comput*, 2020, 32: 1863–1900