A Novel Adaptive Fading Kalman Filter-Based Approach to Time-Varying Brain Spectral/Connectivity Analyses of Event-Related EEG Signals

JIEWEI LI [1,2], **SHING-CHOW CHAN** [2], **(Member, IEEE), ZHONG LIU** [1],
AND CHUNQI CHANG [1], **(Member, IEEE)**

[1] School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen 518060, China
[2] Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong

Corresponding authors: Zhong Liu (liuzhong@szu.edu.cn) and Chunqi Chang (cqchang@szu.edu.cn)

ABSTRACT This paper proposes a novel adaptive fading Kalman filter (AF-KF)-based approach to time-varying brain spectral and functional connectivity analyses of event-related multi-channel electroencephalogram (EEG) signals. By modeling the EEG signals as a time-varying (TV) multivariate autoregressive (MVAR) process, a new AF-KF with variable number of measurements (AF-KF-VNM) is proposed for estimating the spectra of the EEG signals and identifying their functional connectivity. The proposed AF-KF-VNM algorithm uses a new adaptive fading method to adaptively update the model parameters of the KF for improved state estimation and utilizes multiple measurements for better adaptation to the nonstationary signal observations. Experimental results on a simulated data for modeling the TV directed interactions in multivariate neural data show that the proposed AF-KF-VNM method yields better tracking performance than other approaches tested. The proposed algorithm is then integrated into a novel methodology for combined functional Magnetic Resonance Imaging (fMRI) activation maps and EEG spectrum analyses for quantifying the differences in spectrum contents and information flows between the target and standard conditions in a visual oddball paradigm. The results and findings show that the proposed methodology agrees well with the literature and is capable of revealing significant frequency components and information flow involved as well as their time variations.

INDEX TERMS Electroencephalogram (EEG), adaptive fading, Kalman filter (KF), multivariate autoregressive (MVAR), connectivity analysis.

I. INTRODUCTION

There has been a growing interest in developing multivariate approaches for studying brain functions based on non-invasive neuroimaging techniques such as functional Magnetic resonance imaging (fMRI), magnetoencephalography (MEG) and electroencephalography (EEG). Such analyses can provide valuable information regarding the functioning and connectivity of different brain regions for normal as well as diseased patients [1]–[5].

The associate editor coordinating the review of this manuscript and approving it for publication was Pietro Savazzi.

A number of mathematical methods have been proposed to infer the functional and effective connectivity of brains from EEG data [6]. A commonly used approach is based on bivariate method, in which the multivariate connectivity is estimated in a pair-wise manner. Example measures include the Pearson correlation coefficient [7] and coherence function [8], [9]. However, these methods only estimate the correlation between electrodes rather than the direction of the functional links. Therefore, it may yield large number of connections when cortex areas are highly synchronous. To overcome this problem, multivariate methods such as Granger causality [10], partial directed coherence (PDC) [11], and

directed transfer function (DTF) [12], [13], etc., have been proposed. Usually, the multichannel EEG signals is modeled as a multivariate autoregressive (MVAR) model [14] where the inter-channel dependency is modeled by the linear relationship between the current samples with those from the past. Based on the estimated MVAR model, the statistical correlation between the channels at different frequency bands can be inferred. An advantage of the MVAR model is that they can provide directed inference between different bands in different channels. Therefore, it facilitates the interpretation of the information flow related to different neural oscillations involved in the task. Yet another approach is to perform source localization and estimate the directional connectivity from the extracted source signals via source-imaging techniques and connectivity analysis [15], [16]. While it may identify the source signals and their inter-dependency, its performance is highly dependent on the brain propagation model.

In this paper, we propose an alternative methodology by combining the fMRI activation maps and EEG spectra for quantifying the information flows. In particular, we propose an adaptive fading Kalman filter with variable measurement (AF-KF-VNM) -based autoregressive (AR) method to compute the spectra of the EEG signals captured at those electrodes which lie in the close vicinity of the activation maps obtained from the fMRI analysis. This helps to characterize the frequency bands possibly involved in the activated brain regions under the standard and target conditions in a visual oddball experiment. We then compute the information flows between these relevant electrodes using the proposed AF-KF-VNM-based MVAR method and established regions with significant p values for further analysis. Furthermore, a source analysis is performed to validate the findings.

Generally, methods for estimating MVAR model can be classified into two categories, namely 1) the sliding-window approaches, such as short-time directed transfer function (STDTF) [12], [17], and 2) the parametric methods, e.g., Kalman filter (KF)-based algorithms [18]–[22]. The former employs a time shifting window to extract a short-time data segment from the original signal at a time point of interest. These local data samples are assumed to be wide-sense stationarity and hence the MVAR coefficients of each segment can be estimated independently at different time instants using classical AR identification methods [23]. The selection of the window size is crucial to the performance of the sliding-window approach. Too small a window size may yield large variance of the estimated model coefficients, while too large a window size will cause tracking lag that degrades the tracking performance. Therefore, automatic window selection is an important practical problem, which may considerably affect the analysis of nonstationary signals. On the other hand, the parametric method formulates the MVAR model as a state-space model under Gaussian processes. Consequently, the model coefficients can be optimally estimated by the celebrated KF algorithm provided that the model parameters are known a priori. In fact, variants

of KF, called the General Linear Kalman Filter [24]–[26] have been successively applied to analysis of laser-evoked brain potentials and related applications. It treats observations from multiple trials as measurement and hence can be applied to EEG with more channels.[1] A Dual Extended Kalman Filter has also been applied to study the time-varying cortical connectivity analysis of newborn EEG [27]. Moreover, KF has been successively applied in fMRI analysis via Dynamic Causal Modelling (DCM) [28], dynamic granger causality [29], etc. In [29], the KF is applied to infer the causality from the MVAR model of the functional networks obtained by independent component analysis. DCM is based on generative models in which models are compared within a Bayesian framework in order to infer the functional connectivity between neuronal populations or brain regions from the observed EEG, MEG, or fMRI data [28], [29]. In many such real-world applications, the model parameters of the KF are often unknown and they have to be estimated or chosen by trial and error [19], [30], [31]. Therefore, it is desirable to develop new adaptive algorithms for automatic estimation of these model parameters in these KF algorithms.

The proposed AF-KF-VNM algorithm aims to identify the MVAR models from nonstationary EEG signals. It is an extension of our previous KF algorithm for single channel AR model estimation using a variable number of measurements [32], [33] instead of using a single or fixed number of measurements as in the conventional KFs [18]. The use of variable number of measurements enables better adaptation to the nonstationary EEG signals. Moreover, it is found that the conventional KF can be used to implement such KF-VNM algorithm by using measurement and noise covariance matrices with a dimension which varies with the number of measurements used. Furthermore, a novel adaptive fading (AF) method is proposed to adaptively update possible mismatched model parameters in estimating the MVAR coefficients. This helps to alleviate the problem of selecting the unknown or time-varying (TV) model parameters required for the KF. The proposed AF approach differs from the conventional AF KFs [34]–[40] in that the fading parameters are formulated in form of a linear regression, which can be conveniently solved using another low order KF. In addition, the proposed AF approach extends naturally to the KF-VNM case where the measurements dimension is locally time-varying. Due to page limitation, the comparison between the proposed and conventional AF approaches is included in the supplementary material. A comparison between the proposed AF-KF-VNM and the particle filter proposed in [41] is also given. It is shown that that the proposed filter offers significantly better performance than the conventional KF and achieves a comparable performance to that of PF, but at a much lower arithmetic complexity.

---

[1] Our method can also be applied to this setting. On the other hand, we process the data from each trial using our algorithm and obtain its average behavior so as to reduce its sensitive to trial-to-trial variability.

Experimental results on a simulated data for modeling the TV directed interactions in multivariate neural data [42] show that the proposed AF-KF-VNM method yields better performance in tracking the interactions than other approaches tested. Moreover, the estimated time-varying DTF of the proposed AF-KF-VNM algorithm also clearly reveals the influence of the oscillating strength of the various processes.

To illustrate the potential of the proposed methodology in studying TV brain functional connectivity from EEG signals, we focus on real-world multi-channel EEG-fMRI recordings in an oddball paradigm [43]. The results show that the proposed methodology agrees well with the literature and is able to reveal significant frequency components and information flow involved as well as their time variations. These and other findings obtained demonstrate the effectiveness of the proposed AF-KF-VNM-based spectrum/connectivity analyses and the advantages of combined EEG and fMRI analyses to understand the neural mechanism for P300 generation. To our best knowledge, this is the first spectral, connectivity and localization analyses of the EEG-fMRI recordings of the visual oddball paradigm.

The rest of the paper is organized as follows. In Section II, the concept of TV-MVAR model and the DTF metric in measuring connectivity or information flows are described. The proposed AF-KF-VNM algorithm for identification of the MVAR model is given in Section III. Experimental results on both simulated and real oddball data are presented in Section IV. The proposed combined fMRI and EEG analyses and relevant findings are also described. Finally, conclusions are drawn in Section V.

II. TV-MVAR-BASED CONNECTIVITY MEASUREMENT

A. STATIONARY AND TIME-VARYING MULTIVARIATE AR MODELS

In the following, we denote scalars, vectors and matrices by respectively small, bold, and bold-capital letters. In the MVAR modeling of multi-uch as the multi-channel EEG signals, the multi-channel discrete-time signal $s \in R^M$ is modeled as follows:

$$s_k = \sum_{j=1}^{p} A[j]s_{k-j} + v_k, \qquad (1)$$

where $k = 1, 2, \ldots,$ denotes the time index, $s_k = [s_{1,k}, \ldots, s_{M,k}]^T$ with $s_{m,k}, m = 1,\ldots, M$, the signal captured at the m-th channel and time instant k, M denotes the number of channels and p is the model order. The matrices $A[j] \in R^{M \times M}, j = 1, \ldots, p$, correspond to model parameters and $v_k = [v_{1,k}, \ldots, v_{M,k}]^T$ is an independent and identically distributed (i.i.d.) zero mean and temporally white Gaussian noise. That is, $E[v_k] = 0, E[v_k v_k^H] = C_{vv} \succeq 0$, and $E[v_{k'} v_k^H] = 0$, for $k' \neq k$. For stationary processes, $A[j]$ will be independent of time index k, while for time-varying MVAR processes, the matrices $A[j]$ are function of time, i.e. $A_k[j]$.

To estimate the model parameters, one can convert the model in (1) to a conventional linear model using the vectorization operator. Specifically, the matrix parameters can be converted to a vector via the vectorization operator $vec(A) = [a_{1,1}, \ldots, a_{M,1}, a_{1,2}, \ldots, a_{M,2}, \ldots, a_{1,N}, \ldots, a_{M,N}]^T$, which stacks the columns of a matrix $A \in R^{M \times N}$ together and $a_{m,n}$ is the (m, n)- th element of A. Consequently, the MVAR model in (1) can be rewritten in the following vector form:

$$s_k = H_k a + v_k, \qquad (2\text{-}1)$$

where

$$H_k = I_{M \times M} \otimes [s_{k-1}^T, s_{k-2}^T, \ldots, s_{k-p}^T], \qquad (2\text{-}2)$$
$$a = vec((A[1], \ldots, A[p])^T), \qquad (2\text{-}3)$$

$\otimes$ denotes the Kronecker-product of two matrices, and $I_{M \times M}$ represents an $M \times M$ identity matrix. The coefficient matrices $A[j]$ can be readily recovered from the vector a. (2-1) can be solved using the least square method. For time-varying processes, one can imagine that the white noise v_k is passed through a time-varying IIR system in the form

$$s_k = \sum_{l=1}^{p} A_k[l]s_{k-l} + v_k, \qquad (3)$$

where $A[l]$ in (1) is now a time-varying quantity given by $A_k[l]$. The linear model corresponding to (2-1) becomes $s_k = H_k a_k + v_k$, where $a_k = vec((A_k[1], \ldots, A_k[p])^T)$.

Due to the difficulty in estimating the time-varying MVAR coefficients, most conventional methods segment the input signal using the sliding window method and assume that the model coefficients are invariant inside each window. In some cases, the model coefficients are simply assumed to be stationary. Next, we introduce different conventional metrics to measure the connectivity at different frequency bands using the estimated MVAR model.

B. PARAMETRIC TIME-VARYING DTF METRIC

After computing the MVAR model coefficients, the information flow between different channels can be quantitatively estimated from the PDC or DTF measurements. Since DTF captures not only direct, but also cascade flows, between channels, it is adopted here to measure the information flows. For illustration purpose, let's assume for simplicity that v_k is a deterministic signal. By taking the discrete-time Fourier transform (DTFT), $\Im\{.\}$, of (1), one gets:

$$\Im\left\{s_k - \sum_{l=1}^{p} A[l]s_{k-l}\right\} = \Im\{v_k\}$$
$$\Leftrightarrow A(f)s(f) = v(f), \qquad (4)$$

where

$$A(f) = I - \sum_{l=1}^{p} A[l]e^{-j2\pi fl} = \sum_{l=1}^{p} B[l]e^{-j2\pi fl}, \qquad (5)$$

and $\Im\{s_k\} = s(f) = \Sigma_{k=0}^{p} s_k e^{-j2\pi fk}$ and $\Im\{v_k\} = v(f) = \Sigma_{k=0}^{p} v_k e^{-j2\pi fk}$ are the DTFT of s_k and v_k, respectively. The frequency range is $f \in [-f_s/2, f_s/2]$ where f_s is the sampling frequency, and $j = \sqrt{-1}$. Eq. (4) can also be written as $s(f) = (A(f))^{-1}v(f) = G(f)v(f)$, where $G(f)$ is the matrix transfer function at the k-*th* time instant. If $v(f)$ is a temporally white random process, then the DTFT of v_k and s_k may not exist but one can work with the autocorrelation of s_k and v_k and the power spectral density of s_k is then given by $S_s(f) = (A(f))^{-1}(A(f))^{-H} S_v$, where $S_v = E[v_k v_k^H] = C_v$ is the power spectral density of v_k, which is a positive semidefinite matrix. Moreover, if one assumes that the process is slowly time-varying, then the DTFT of the locally windowed samples or short-time DTFT can be written approximately as $s_k(f) = (A_k(f))^{-1}v(f) = G_k(f)v(f)$, where $G_k(f)$ is the matrix transfer function at the k-*th* time instant. The power spectral density of s_k is then given by $S_{s_k}(f) = (A_k(f))^{-1}(A_k(f))^{-H} S_v$.

Many measures have been proposed to measure the functional connectivity from the matrix transfer function $G(f)$. For instance, the following normalized DTF [12] can be used to describe the information flow from channel l to channel m

$$c_{m,l}(f) = \frac{|G_{m,l}(f)|}{\sqrt{\sum_{l=1}^{M} |G_{m,l}(f)|^2}}, \tag{6}$$

where $G_{m,l}(f)$ is the (m, l)-th element of $G(f)$. It can be seen that $c_{m,l}(f)$ is the normalized absolute value of $G_{m,l}(f)$, which has a value between 0 and 1, i.e. [0, 1]. Similarly, the following time-varying normalized DTF can be used to describe the information flow from channel l to channel m at time instant k:

$$c_{m,l}(k,f) = \frac{|G_{m,l}(k,f)|}{\sqrt{\sum_{l=1}^{M} |G_{m,l}(k,f)|^2}}, \tag{7}$$

where $G_{m,l}(k,f)$ is the (m, l)-th element of $G_k(f)$. As mentioned, due to the difficulty in estimating the time-varying MVAR coefficients, most conventional methods employ the sliding window method and assume that the model coefficients are invariant inside each window. Consequently, the performance is highly dependent on the selection of window size. We now introduce the proposed AF-KF-VNM algorithm for estimating the time-varying MVAR model and hence the connectivity from multichannel EEG signals.

III. TV-MVAR-BASED CONNECTIVITY MEASUREMENT
In this section, we first formulate the problem of estimating the TV MVAR model parameters as a state-space model and introduce the celebrated KF for tracking the model parameters. We then introduce the KF-VNM algorithm which utilizes a variable number of measurements (VNM) for better identification of the TV coefficients in nonstationary scenarios. Finally, a new adaptive fading (AF) method

is proposed for adaptively updating the covariance matrices of the KF-VNM algorithm, which yields the proposed AF-KF-VNM algorithm.

A. KALMAN FILTER FOR MVAR MODELS
One of the difficulties in estimating the TV MVAR model in (3) is to determine the size of the window, inside which the MVAR model is to be estimated. Too large a window size will introduce significant bias in the estimated MVAR parameters, whereas too small a window size will introduce significant variance in the estimated parameters. Moreover, it is important to ensure the continuity of the estimated parameters. In fact, each sample of (2-1) in the neighborhood of the current time instant can be viewed as the observation or data term in conventional maximum likelihood function whereas the continuity can be viewed as the prior information in the maximum a posterior probability estimation. An effective recursive method to incorporate such prior information in (2-1) is to introduce an additional dynamic equation, which leads to the following state-space model:

$$a_k = F_k a_{k-1} + w_k, \tag{8-1}$$
$$s_k = H_k a_k + v_k, \tag{8-2}$$

where F_k is the state transition matrix, w_k and v_k are respectively the state and measurement noises which are assumed to be mutually independent and zero mean Gaussian distributed random vector, i.e. $w_k \sim N(0, Q_k)$ and $v_k \sim N(0, R_k)$. Normally, F_k is chosen as an identity matrix $I_{M^2p \times M^2p}$, which gives rise to the so-called random walk model. (8-1) implies that a_k is equal to a_{k-1} except for an additional random error distributed as $N(0, Q_k)$. This helps to enforce the continuity constraint. An advantage of such state space formulation is that the state variable a_k can be optimally estimated in the minimum mean square error sense using the celebrated KF algorithm. More precisely, with appropriate initialization of the state vector $\hat{a}_{0|0}$ and its covariance $P_{0|0}$, the KF estimates a_k using the following prediction and correlation steps:

Prediction

$$\hat{a}_{k|k-1} = F_k \hat{a}_{k-1|k-1}, \tag{9-1}$$
$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k, \tag{9-2}$$

Correction

$$e_k = s_k - H_k \hat{a}_{k|k-1}, \tag{9-3}$$
$$K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1}, \tag{9-4}$$
$$\hat{a}_{k|k} = \hat{a}_{k|k-1} + K_k e_k, \tag{9-5}$$
$$P_{k|k} = (I - K_k H_k) P_{k|k-1}, \tag{9-6}$$

where e_k is the prediction error of the measurement vector, K_k is the Kalman gain matrix, and $\hat{a}_{k|k-1}$ and $\hat{a}_{k|k}$ represent respectively the estimator of a_k given the measurements up to time instant $k-1$ and k. $P_{k|k-1}$ and $P_{k|k}$ are the corresponding covariance matrices of $\hat{a}_{k|k-1}$ and $\hat{a}_{k|k}$.

B. KF WITH VARIABLE NUMBER OF MEASUREMENTS (KF-VNM)

Though (8) can be estimated optimally using the KF if the state and measurement noise parameters are known, it only makes use of the current measurement and relies entirely on the state dynamic to reduce the variance of the estimation. Like the sliding window approach, it is advantageous to employ as many as possible neighboring samples to reduce the estimation variance. However, care has to be taken to select a proper number of measurements. It is because the use of more measurement will help to reduce the bias error due to the prior if the system state is fast time-varying. On the other hand, when the state changes slowly, including more measurements can help to reduce the estimation variance. Ideally, a variable number of measurements should be employed in the measurement equation. Suppose that the set of measurements used for state estimation at time instant k lies in a symmetric window around s_k, $\bar{s}_k = [s_{k-W_k}^T, s_{k-W_k+1}^T \ldots, s_k^T, \ldots, s_{k+W_k-1}^T, s_{k+W_k}^T]^T$, where $(2W_k + 1)$ is the window size. Including the augmented measurement vector into (8-2) yields:

$$\bar{s}_k = \overline{H}_k a_k + \bar{v}_k, \quad (10)$$

where $\overline{H}_k = [H_{k-W_k}^T, H_{k-W_k+1}^T, \ldots H_k^T, \ldots, H_{k+W_k-1}^T, H_{k+W_k}^T]^T$ and the measurement noise vector becomes $\bar{v}_k \sim N(0, \ \overline{R}_k)$ with $\tilde{R}_k = blkdiag\{R_{k-W_k}, R_{k-W_k+1}, \ldots, R_k, \ldots, R_{k+W_k-1}, R_{k+W_k}\}$. Here, $blkdiag\{\cdot\}$ denotes the diagonal concatenation of block matrices. The variable measurement equations in (10) can be augmented to the state equation in (8-1) to form a measurement-augmented state-space model at each time instant. Moreover, we can see that the previously described KF algorithm in Section III-A is also applicable to this augmented system by replacing H_k and R_k in (9-4) with $\overline{H}_k$ and $\overline{R}_k$ respectively. Note that the window size $2(W_k + 1)$ and hence $\overline{H}_k$ and $\overline{R}_k$ will vary according to the signal at each iteration. In the context of EEG signal analysis, slight delay can usually be tolerated and hence a symmetric window can be used. If such delay is undesirable, a non-symmetric or even a one-sided window can be used to reduce the delay due to the use of future measurements.

Next, we shall introduce a method for determining the number of elements to be used in the augmented measurement vector. It is based on the observation that significant state changes are usually associated with fast varying parameters and the number of neighboring measurements to be used should be reduced accordingly. To this end, we estimate the derivative or change of the system state using the following equations:

$$c_k = \hat{a}_{k-1} - \tilde{a}_{k-1}, \quad (11\text{-}1)$$

$$\tilde{a}_k = \lambda \tilde{a}_{k-1} + (1 - \lambda)\hat{a}_{k-1}, \quad (11\text{-}2)$$

where c_k is the approximated time derivative of the state estimate $\hat{a}_k$, and λ is a positive forgetting factor less than one for smoothing $\hat{a}_k$ to form $\tilde{a}_k$. The change can then be measured using the 1-norm of c_k, which is denoted by $\|c_k\|_1$.

When the TVAR process varies slowly, $\|c_k\|_1$ will decrease and converge gradually from its initial value to a very small value. This idea has been used previously in [22] for varying the step-size in the LMS-type algorithms. In contrast, $\|c_k\|_1$ will become large while the system changes rapidly. Therefore, $\|c_k\|_1$ serves as a measure of the time variation of the state, and it can be used to determine the window size. More specifically, the following four-step algorithm is proposed to determine the number of elements in the augmented measurement vector $\bar{s}_k$:

Step 1. Compute the absolute value of the approximate derivative of $\|c_k\|_1$ as: $\beta_k = \big| \|c_k\|_1 - \|c_{k-1}\|_1 \big|$.

Step 2. Smooth β_k by averaging it over a time window of length N_s to obtain: $\tilde{\beta}_k = \frac{1}{N_s} \Sigma_{k=1}^{N_s} \beta_k$.

Step 3. Normalize $\tilde{\beta}_k$ with the first N_0 data as follows: $\overline{\beta}_k = \frac{1}{N_s \tilde{\beta}_0} \Sigma_{k=1}^{N_s} \beta_k$, where $\tilde{\beta}_0 = \frac{1}{N_0} \Sigma_{k=1}^{N_0} \beta_k$.

Step 4. Calculate the number of elements as follows: $W_k = \max\{L_L, round[L_L + (1 - \overline{\beta}_k)(L_U - L_L)]\}$, where L_L and L_U denotes respectively the lower and upper bounds of W_k and $L_U \geq L_L \geq 0$.

It can be seen from the equation in step 4 that as $\overline{\beta}_k$ approaches 1, which indicates that the estimated state variation is getting large, W_k will approach the lower bound L_L. This means that a smaller number of measurements will be chosen. When the estimated state change is small, $\overline{\beta}_k$ will be nearly zero and a large W_k will be used. This yields the proposed KF-VNM algorithm, which uses a variable number of $\overline{W}_k = 2W_k + 1$ measurements to estimate the system state recursively.

C. ADAPTIVE FADING KF WITH VNM (AF-KF-VNM)

The conventional KF algorithm and the KF-VNM above assume that the state and measurement noise covariance are known a priori. However, this information may not be available and hence its performance may be degraded, especially for time-varying or environmental dependent scenarios. For better adaptability and hence estimation of the states, we propose in this section a new AF-KF-VNM algorithm based on the adaptive fading method for updating these parameters.

The basic idea is to update the nominal covariance of the state and measurement noises by computing their empirical averages. Though there are previous works on using adaptive fading [34]–[40], to our best knowledge, the proposed approach is the first that can deal with a variable number of measurements in a single framework. Also, the proposed AF approach has a better performance than these conventional algorithms as shown in the supplementary material. Specifically, the prediction error in the KF-VNM can be calculated by modifying (9-3) as follows

$$e_k = \bar{s}_k - \overline{H}_k \hat{a}_{k|k-1}, \quad (12)$$

from which, one can obtain the following equation

$$P_k^e = E(e_k e_k^T) = \overline{H}_k \tilde{P}_{k|k-1} \overline{H}_k^T + \tilde{R}_k, \quad (13)$$

relating the covariance of the prediction error with the true state error covariance $\tilde{P}_{k|k-1}$ and the measurement noise covariance $\tilde{R}_k$. They may be different from their nominal values which are denoted by $P_{k|k-1}$ and $\overline{R}_k$, respectively. Here, we aim to find positive scalar correction or fading factors $f_{s,k}$ and $f_{m,k}$ for $P_{k|k-1}$ and R_k respectively such that

$$\tilde{P}_{k|k-1} = f_{s,k} P_{k|k-1}, \tag{14}$$
$$\tilde{R}_k = f_{m,k} \overline{R}_k, \tag{15}$$

so as to satisfy (13). Substituting (14) and (15) into (13) gives an equation for the desired factors in terms of other known quantities:

$$P_k^e = f_{s,k} \overline{H}_k P_{k|k-1} \overline{H}_k^T + f_{m,k} \overline{R}_k. \tag{16}$$

Instead of matching the whole matrix, we take the trace of the left and right-hand sides of (16) and it gives the following equation: $tr(P_k^e) = f_{s,k} tr(H_k P_{k|k-1} H_k^T) + f_{m,k} tr(\overline{R}_k)$. Here, the measurement residual covariance P_{k^e} can be estimated empirically as

$$\hat{P}_k^e = \frac{\overline{W}_k}{M} \sum_{t=k-M+1}^{k} \frac{e_t e_t^T}{\overline{W}_t}, \tag{17}$$

where M specifies the window size for estimating the empirical covariance. Consequently, the estimation of the fading factors reduces to the following recursive linear regression problem over time:

$$tr(\hat{P}_k^e) \approx f_{s,k} tr(\overline{H}_k P_{k|k-1} \overline{H}_k^T) + f_{m,k} tr(\overline{R}_k), \tag{18}$$

where $k = 1, 2\ldots, f_{s,U} \geq f_{s,k} \geq f_{s,L} \geq 0, f_{m,U} \geq f_{m,k} \geq f_{m,L} \geq 0$, and $f_{s,U}$ and $f_{s,L}$ are respectively the upper and lower bounds of $f_{s,k}$ (likewise for $f_{m,k}$). The purpose of the normalization factor $(\overline{W}_t)^{-1}$ in (17) is to ensure that each equation in (18) over time will have the same influence, as the number of measurements used at each time instant $\overline{W}_k$ may be different.

To solve this recursive linear regression problem, one can use a KF with the following state-space model:

$$x_k^c = F_k^c x_{k-1}^c + \xi_k, \xi_k \sim N(0, Q_k^c), \tag{19-1}$$
$$y_k^c = H_k^c x_k^c + \eta_k, \eta_k \sim N(0, R_k^c), \tag{19-2}$$

where $x_k^c = [f_{s,k} \quad f_{m,k}]^T$ is the state vector containing the fading factors, $H_k^c = [tr(\overline{H}_k P_{k|k-1} \overline{H}_k^T), \quad tr(\overline{R}_k)], F_k^c = I, y_k^c = tr(\hat{P}_k^e)$, and ξ_k and η_k are the corresponding state and measurement noises, respectively. This allows the fading factors to be estimated recursively using the conventional KF with a random walk model. Moreover, during each recursion, the estimated parameters will be projected to the lower and upper bound constraints above. Finally, we summarize the proposed AF-KF-VNM algorithm in Table 1.

TABLE 1. The proposed AF-KF-VNM algorithm.

S1: **Initialization:** $\hat{x}_{0|0}$, $P_{0|0}$
for $k = 1, 2, \ldots N$
S2: **Prediction:**

$\quad \hat{a}_{k|k-1} = F_k \hat{a}_{k-1|k-1}$,

$\quad P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k$.

S3: **Estimate** W_k **using four-step algorithm and form:**

$\quad \overline{s}_k = [s_{k-W_k}^T, s_{k-W_k+1}^T \cdots, s_k^T, \cdots, s_{k+W_k-1}^T, s_{k+W_k}^T]^T$

$\quad \overline{H}_k = [H_{k-W_k}^T, H_{k-W_k+1}^T, \cdots H_k^T, \cdots, H_{k+W_k-1}^T, H_{k+W_k}^T]^T$,

$\quad \overline{R}_k = blkdiag(R_{k-W_k}, R_{k-W_k+1} \cdots, R_k, \cdots, R_{k+W_k-1}, R_{k+W_k})$.

S4: **Covariance compensation:**

$\quad e_k = \overline{s}_k - \overline{H}_k \hat{a}_{k|k-1}$,

$\quad \hat{P}_k^e = \dfrac{\overline{W}_k}{M} \displaystyle\sum_{t=k-M+1}^{k} \dfrac{e_t e_t^T}{\overline{W}_t}$.

$\quad$ Compute $f_{s,k}$ and $f_{m,k}$ by solving:

$\quad tr(\hat{P}_k^e) \approx f_{s,k} tr(\overline{H}_k P_{k|k-1} \overline{H}_k^T) + f_{m,k} tr(\overline{R}_k)$.

$\quad$ Correct the covariance:

$\quad P_{k|k-1} \leftarrow f_{s,k} P_{k|k-1}$,

$\quad \overline{R}_k \leftarrow f_{m,k} \overline{R}_k$.

S5: **Correction:**

$\quad K_k = P_{k|k-1} \overline{H}_k^T (\overline{H}_k P_{k|k-1} \overline{H}_k^T + \overline{R}_k)^{-1}$,

$\quad \hat{a}_{k|k} = \hat{a}_{k|k-1} + K_k e_k$,

$\quad P_{k|k} = (I - K_k \overline{H}_k) P_{k|k-1}$.

end for k

Note: "$\leftarrow$" denotes a simple replacement of the left with the right.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

We first test the AF-KF-VNM method and compare it with other conventional methods (including recursive-least square (RLS), RLS with a constant forgetting factor (RLS-FF), the conventional KF [18], Adaptive AR model (AAR) [19] and short time DTF (STDTF) [12], [17]) for identifying simulated MVAR models. Moreover, to examine the effectiveness of using future data for estimation, we also implement an online version of the proposed AF-KF-VNM, AF-KF-VNM-ONLINE, which is the AF-KF-VNM with one-sided window containing only past and the current measurements and compare it with the AF-KF-VNM using symmetric window and the Kalman smoother (KS). Then the proposed AF-KF-VNM algorithm is used to study the brain connectivity from a real-world EEG recording in an oddball paradigm.

A. EVALUATION ON SIMULATION DATA

1) SIMULATION MODEL DESCRIPTION

We first evaluate the performance of the various algorithms using a simulated data set consisting of a 3-dimensional MVAR process. The data has been previously used to assess the time-varying directed interactions in multivariate neural data [42], which contains a stochastically driven relaxator and two damped stochastically driven oscillators with
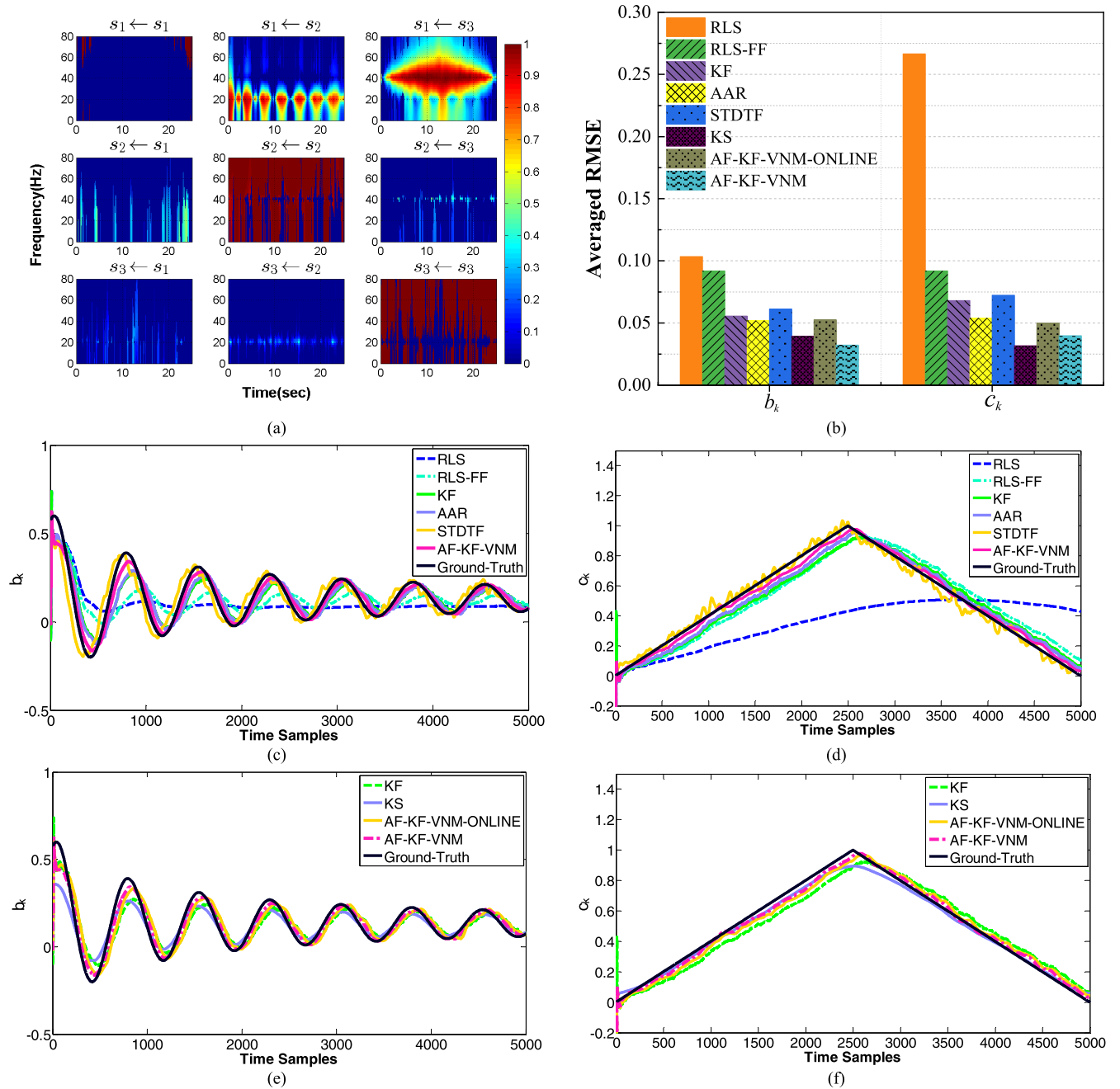
FIGURE 1. Simulation results of MVAR(2): (a) the computed DTF by using the proposed AF-KF-VNM algorithm; (b) summarizes the averaged RMSEs during the estimation of b_k and c_k using various algorithms; (c)-(f) plot the model parameter b_k and c_k obtained by the tested algorithms.

time-varying interaction as follows:

$$s_{1,k} = 0.59s_{1,k-1} - 0.2s_{1,k-2} + b_k s_{2,k-1} + c_k s_{3,k-1} + v_{1,k},$$
$$s_{2,k} = 1.58s_{2,k-1} - 0.96s_{2,k-2} + v_{2,k},$$
$$s_{3,k} = 0.60s_{3,k-1} - 0.91s_{3,k-2} + v_{3,k}. \tag{20}$$

The time-varying parameters b_k and c_k are plotted in Fig. 1 with black lines, which serve as ground truth for subsequent comparisons. It can be seen that the process $s_{1,k}$ is influenced by $s_{2,k}$ via b_k as oscillating strength, whereas $s_{3,k}$ influence $s_{1,k}$ with increasing strength in the first half

of the simulation and with decreasing strength in the second half.

2) PARAMETER SETTINGS

The specific configurations of various algorithms are summarized as follows: 1) For the RLS-FF algorithm, the constant forgetting factor used is 0.996. 2) For the KF, the covariance of the state and measurement noises are respectively 10^{-5} and 1, which were manually chosen by trial and error. 3) For the AAR algorithm, the update coefficient is set as 0.003

and the state noise covariance is adaptively calculated based on the trace of the Kalman gain matrix. 4) In the STDTF, the window length is chosen as 100 with 99 overlap. The least square method is used to estimate the model parameters in each window. 5) In the proposed AF-KF-VNM algorithm, the settings for the fading factors are $1.001 \geq f_{s,k} \geq 1$, $1.001 \geq f_{m,k} \geq 1$, $Q_k^c = diag(10^4, 10^{-4})$ and $R_k^c = 1$. The settings for the KF part in the AF-KF-VNM are the same as in 2) for a fair comparison. To determine the number of measurements, we use $\lambda = 0.999$, $N_s = N_0 = 50$, $L_L = 50$ and $L_U = 100$. 6) The configuration for AF-KF-VNM-ONLINE is identical to that of AF-KF-VNM, except that no future measurements is used in order to mimic the case for online estimation without any sample delay. 7) For the KS, the concept of Rauch–Tung–Striebel backward smoothing is employed to update the results of KF using the whole measurements.

3) SIMULATION RESULTS
The estimated time-varying DTF of the proposed AF-KF-VNM algorithm is shown in Fig. 1a, where we can see that the influence of the oscillating strength of process $s_{2,k}$ on $s_{1,k}$ is correctly identified. We can also see that the increasing and decreasing influence of $s_{3,k}$ on $s_{1,k}$ is clearly revealed. The performance of the proposed algorithm is compared with those from the other algorithms (including RLS, RLS-FF, KF, AAR and STDTF) in Figs. 1c and 1d. The plotted b_k and c_k parameters are averaged by 10 Monte Carlo realizations. It can be seen from Figs. 1b and 1c that, among all algorithms tested, the RLS algorithm (with a forgetting factor of one) gives the worst performance as all past data up to the current time instance are used in the estimation. The performance of RLS-FF, KF and AAR algorithms are quite similar and much better than that of the RLS algorithm. However, all of them exhibit tracking lags in estimating the model coefficients. In contrast, the STDTF and AF-KF-VNM algorithms offer improved tracking capacity without incurring significant delays as compared with the other algorithms. Such improvement can be explained by the fact that both STDTF and the proposed AF-KF-VNM methods have utilized future data in their estimation, which help to mitigate the delay caused by estimation from past observations alone. Moreover, as compared with the STDTF method, the AF-KF-VNM algorithm offers much better performance in terms of a smaller tracking variance and more accurate tracking results. This demonstrates the effectiveness of the adaptive fading approach in the proposed AF-KF-VNM algorithm for updating the covariance matrices.

To examine the effectiveness of using future data for current estimation, we also compare the AF-KF-VNM with symmetric window against the online version of AF-KF-VNM (AF-KF-VNM-ONLINE) and the Kalman smoother (KS). The estimated parameters of b_k and c_k are plotted in Fig. 1e and It can be seen that there are more time delays for the estimates obtained by AF-KF-VNM-ONLINE in comparison with those of AF-KF-VNM using symmetric window. This suggests that the use of future data in the AF-KF-VNM

algorithm can alleviate the lagging problem to some extent. Similar phenomenon can be observed by comparing the results between KF and KS in Fig. 1e and 1f, where it can be seen that the lagging problem is significantly mitigated by the KS which uses the entire measurements for state estimation. However, smoothing with the entire data as done in the KS may result in over-smoothed estimates. This is evidenced by the inferior estimates of KS around the peak point at time sample 2500 in Fig. 1f. In contrast, the AF-KF-VNM algorithm using symmetric window is able to offer better estimates which is attributed to the use of variable number of measurements in the AF-KF-VNM method. Finally, the performance of tested algorithms for the estimation of b_k and c_k are compared in terms of the following averaged root mean squared error (RMSE):

$$RMSE_{AV} = \frac{1}{N} \sum_{k=1}^{N} \left(\sqrt{\frac{1}{M} \sum_{m=1}^{M} (x_k(m) - \hat{x}_k(m))^2} \right),$$
(21)

where $x_k(m)$ and $\hat{x}_k(m)$ denote respectively the true and estimated parameters at the m-th Monte Carlo run. The quantitative results are summarized in Fig. 1b. It can be seen that the mean estimation accuracy of the AF-KF-VNM algorithm and the KS are significantly better than those obtained by the other approaches.

B. ANALYSIS OF fMRI/EEG SIGNALS IN THE ODDBALL EXPERIMENT
As mentioned earlier, the studies of the oddball paradigm is inspired by the P300 wave, an ERP component elicited during decision making which is also utilized in brain-computer interface [44], [45]. Previous studies [46] have shown that P300 components can be maximally observed from parietal electrodes. Its generation is also known to be related to attention and memory processing. This suggests that there may be considerable information flows between parietal and other brain areas, which we shall now analyze.

1) DATA DESCRIPTION AND EXPERIMENTAL SETTINGS
We first evaluate the performance of the various algorithms on a real-world oddball experiment from the OpenfMRI database [43] with accession ID ds000116. Seventeen subjects participated in the visual oddball paradigm. During the oddball experiment, subjects were instructed to respond to target stimuli and ignore standard stimuli, which are displayed sequentially in the center of the screen. Main variables manipulated in the paradigm are stimulus type and frequency. In the experiment, the target and standard stimulus are denoted by a large red circle and a small green circle respectively. The target and standard stimuli appear alternately with a 2–3s uniformly distributed variable inter-trial interval (ITI) and each stimulus lasts for 200 milliseconds (ms). There are totally 375 stimuli in the oddball experiment, 20% of which are target. When target stimulus appears in sight, participants are required to press on an MR-compatible

button response pad with the right index finger. Scalp EEG and functional magnetic resonance image (fMRI) data are simultaneously recorded during the experiment. Functional echo-planar image (EPI) data are continuously collected by a 3T Philips Achieva MRI scanner. The whole brain data are covered by obtaining 32 slices of 64×64 voxels with a 2000 ms repetition time (TR) and 25 ms echo time (TE). EEG data are simultaneously recorded using a custom-built MR-compatible EEG system and 43 bipolar pairs are recorded. Before EEG analysis, 43 bipolar EEG signals are re-referenced to the 32-channel electrode space.

2) METHODOLOGY

To obtain a better picture of the brain activities involved in the experiment, we first identify the regions with significant difference ($p < 0.001$) in activation between the standard and target stimulating conditions in the fMRI data using the SPM toolbox (details are given later in this section). We then employ the AF-KF-VNM-based AR model to compute the spectra of the signals captured at those electrodes which lie in the close vicinity of the activation maps obtained from the fMRI analysis. This helps to characterize the significant frequency bands possibly involved in the activated brain regions under the two conditions. We then computed the information flows between these relevant electrodes using the proposed AF-KF-VNM-based MVAR method and established regions with significant p values. The combined analyses allow us to contrast the agreement of the two analyses as well as revealing more information regarding the spectrum composition of the EEG signals involved their time variations and possible information flow.

fMRI Data Processing: The processing of fMRI data follows the standard procedure and is based on the SPM toolbox with first-level statistics including slice-time correction, spatial realignment, spatial normalization, and smoothing [47]. The regressors for first-level statistics are derived from the timing of stimulus onsets convolved with a modeled hemodynamic response. Then, second-level group analysis is performed across two conditions through 17 subjects. The fMRI activation maps are shown in Fig. 2a.

EEG Data Processing: Since the gradient artifacts and ballistocardiogram (BCG) artifacts in the dataset [48] were removed by the authors, the following preprocessing procedure of the gradient-free EEG data in the oddball paradigm is adopted: the averaged reference signal (obtained by averaging the data of all electrodes) is firstly subtracted from the data. Then, the resulting signals from -300ms to 1000ms post-stimulus are extracted and further corrected with a baseline reference [49] calculated using data from -300ms to 0ms. Finally, artifact-free trials are down-sampled to 256 Hz for subsequent estimation of information flows between different channels. We also discarded trials containing eye-blink artifacts by visual inspection, and about 95% of the target trials are retained for analysis.
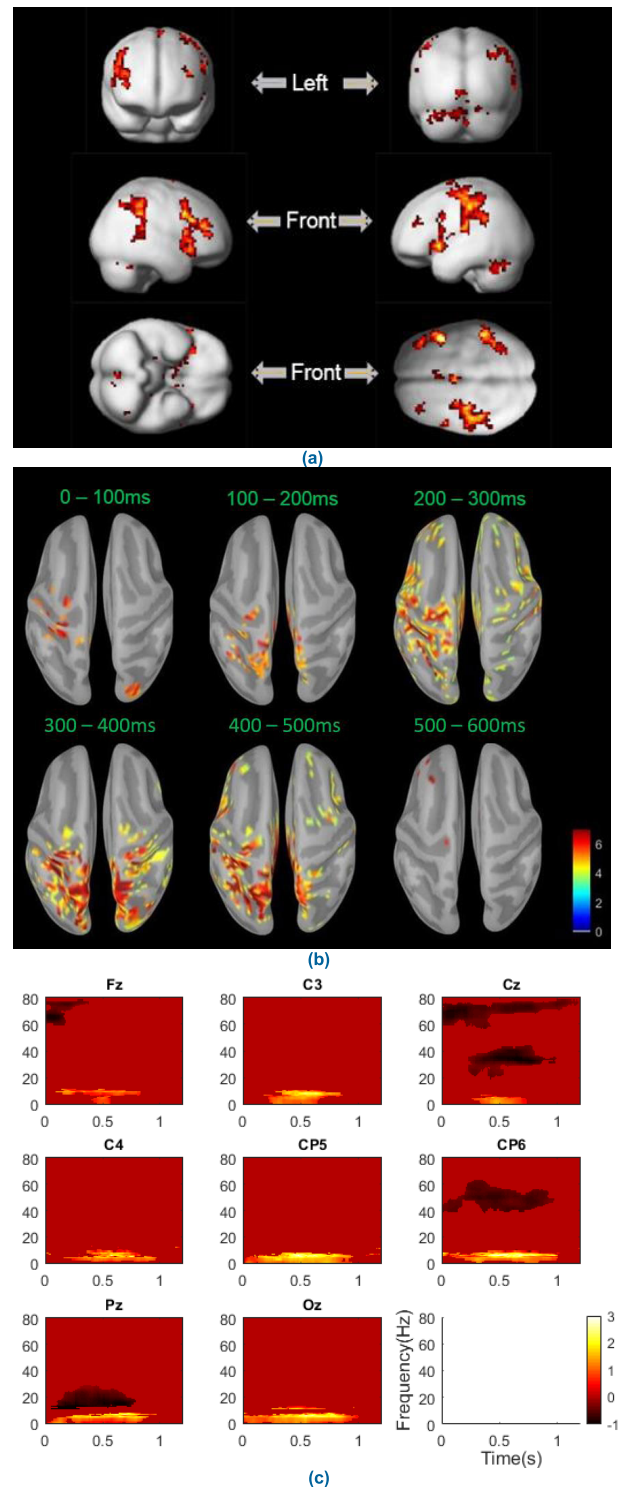

FIGURE 2. (a) fMRI activation maps. Group statistical maps of blood oxygenation level-dependent signal difference for the target condition compared with standard condition, p<0.001 (in red color). (b) Source difference between target and standard conditions in different time durations (p<0.01). The color bar means the absolute mean of the target condition minus that of the standard condition. (c) Power spectrum differences between target and standard conditions. Spectrum power in target condition is significantly higher and lower than standard condition in yellow and black areas, respectively (p<0.05).

Since previous studies [46], [50], [51] have found that the absolute power topography of the ERP response is mainly concentrated in the frontal, central and parietal areas, and according to our fMRI activation analysis in Fig. 2a, eight relevant channels, namely Fz, C3, Cz, C4, CP5, CP6, Pz and Oz, are selected for subsequent spectrum and brain connectivity analysis. The locations of these electrodes in the standard EEG system are shown in the supplementary materials. Frequency range below 80 Hz is used for assessment.

For the spectrum analysis of the eight relevant channels, we model the variation of the signal of each channel by an AR process, the coefficients of which are estimated using the proposed AF-KF-VNM algorithm. The spectra can then be gauged from the estimated AR coefficients as in [32], [41]. The settings of the AF-KF-VNM-based AR method are as follows: 1) the AR model order is selected using the approach in [52] by evaluating the model order on randomly sampled EEG trails in terms of Bayesian information criterion (BIC) [53]. In our experiment, the orders of the randomly sampled trails have a mean value of 7.89 with a standard deviation of 1.07. Thus, the model order is set to 8. 2) The parameters for fading factor estimation are set: $1.02 \geq f_{s,k} \geq 1$, $1.001 \geq f_{m,k} \geq 1$, $\boldsymbol{Q}_k^c = diag(10^4, 10^{-4})$ and $\boldsymbol{R}_k^c = 1$. 3) The covariance of the state and measurement noises of the KF in the AF-KF-VNM algorithm are 10^{-5} and 1, respectively. 4) For the calculation of the number of measurements, we set $\lambda = 0.999$, $N_s = N_0 = 50$, $L_L = 50$ and $L_U = 100$. Compared to the settings of the AF-KF-VNM in the simulation above, most of the parameters are kept unchanged except that the lower and upper bounds in estimating the fading factors are slightly adjusted according to the dynamic characteristic of the signals.

For the brain functional connectivity analysis, the information flows between different electrodes are measured by the normalized DTF metric where the MVAR coefficients are estimated using various approaches including AAR, STDTF and the proposed AF-KF-VNM algorithms for comparison. The settings for different algorithms are as follows: 1) For the AAR algorithm, the update coefficient is set as 0.008, which yields a comparable performance to our best efforts in running the algorithm multiple times by manually tuning the coefficient. 2) For the STDTF, we also used the least square method to estimate the model parameters. Two different window lengths, 50 with 49 overlap and 100 with 99 overlap, are used to generate the results for comparison. 3) For the AF-KF-VNM algorithms, the parameters for fading factor estimation are: $1.001 \geq f_{s,k} \geq 1$, $1.001 \geq f_{m,k} \geq 1$, $\boldsymbol{Q}_k^c = diag(10^4, 10^{-4})$ and $\boldsymbol{R}_k^c = 1$; The covariance of the state and measurement noises of the KF in the AF-KF-VNM algorithm are 10^{-5} and 1, respectively; For the calculation of the number of measurements, we set $\lambda = 0.999$, $N_s = N_0 = 50$, $L_L = 100$ and $L_U = 200$. Comparing with the setting in the simulation, most of the parameters remain the same except the lower and upper limits are slightly adjusted to allow for better adaptation in terms of the number of measurement used. 4) The model order p in the MVAR model for all algorithms is automatically determined by the Bayesian Information Criterion (BIC).

3) EXPERIMENTAL RESULTS AND ANALYSIS
a: fMRI DATA ANALYSIS
Figure 2a depicts the fMRI activity in the target condition versus the standard condition. Regions with significant changes in activity generated by the target stimuli ($p < 0.001$) are shown in red. They include the thalamus, bilateral inferior parietal lobules, precentral sulcus (PrCS), left postcentral gyrus, prefrontal cortex (PFC) and right middle frontal gyri. They were also identified in previous fMRI studies [47], [51], [54], [55]. Although visual P300 is highly activated around channel Pz in EEG study, no significant difference was found in the middle parietal area between the two conditions in the fMRI activation map. The activated locations are summarized in the supplementary material.

b: EEG DATA ANALYSIS
Spectrum Analysis:
To facilitate the spectral analysis, the mean spectrum values of baseline (from -0.5s to 0s) for each frequency bin were removed from the calculated spectrum matrix. Then, a cluster-based permutation approach [56] was used to estimate the spectral difference between two conditions with 2000 permutations for the correction of multiple comparisons.

The significant difference ($p<0.05$) in the spectrum between the target and standard conditions are shown in Fig. 2c for the eight channels chosen. From the results, we find that the target condition has higher low-frequency power (under 12Hz) and lower high-frequency power (over 15Hz) than the standard condition. The main differences are observed from 200ms to 800ms after stimulus onset. For example, in the target condition, significantly increased power under 10Hz during 200ms to 800ms is observed at channel Pz. A decreased power between 15Hz to 40Hz is also observed during the same period. This result is also consistent with the previous study [52].

Functional Connectivity Analysis:
To facilitate the analysis of the information flow between different functional regions of the brain, baseline connectivity patterns were firstly removed from the calculated information flows. This is completed by subtracting the averaged DTF metric within the baseline period from the DTF metric for each trial in both the target and standard task conditions. Then, the differences of the connectivity are obtained by subtracting the DTF metric in standard condition from that in the target, which are referred to as effective information flow. Finally, the independent samples t-test (2 tails) was performed to measure the significance of the functional connectivity. The results obtained with various algorithms, including AAR, STDTF and the proposed AF-KF-VNM algorithms, are presented in Fig. 3 for visual assessment. Regions with significant differences in activity generated by the target stimulus (*p*-values < 0.05) are marked in boxes with black color.
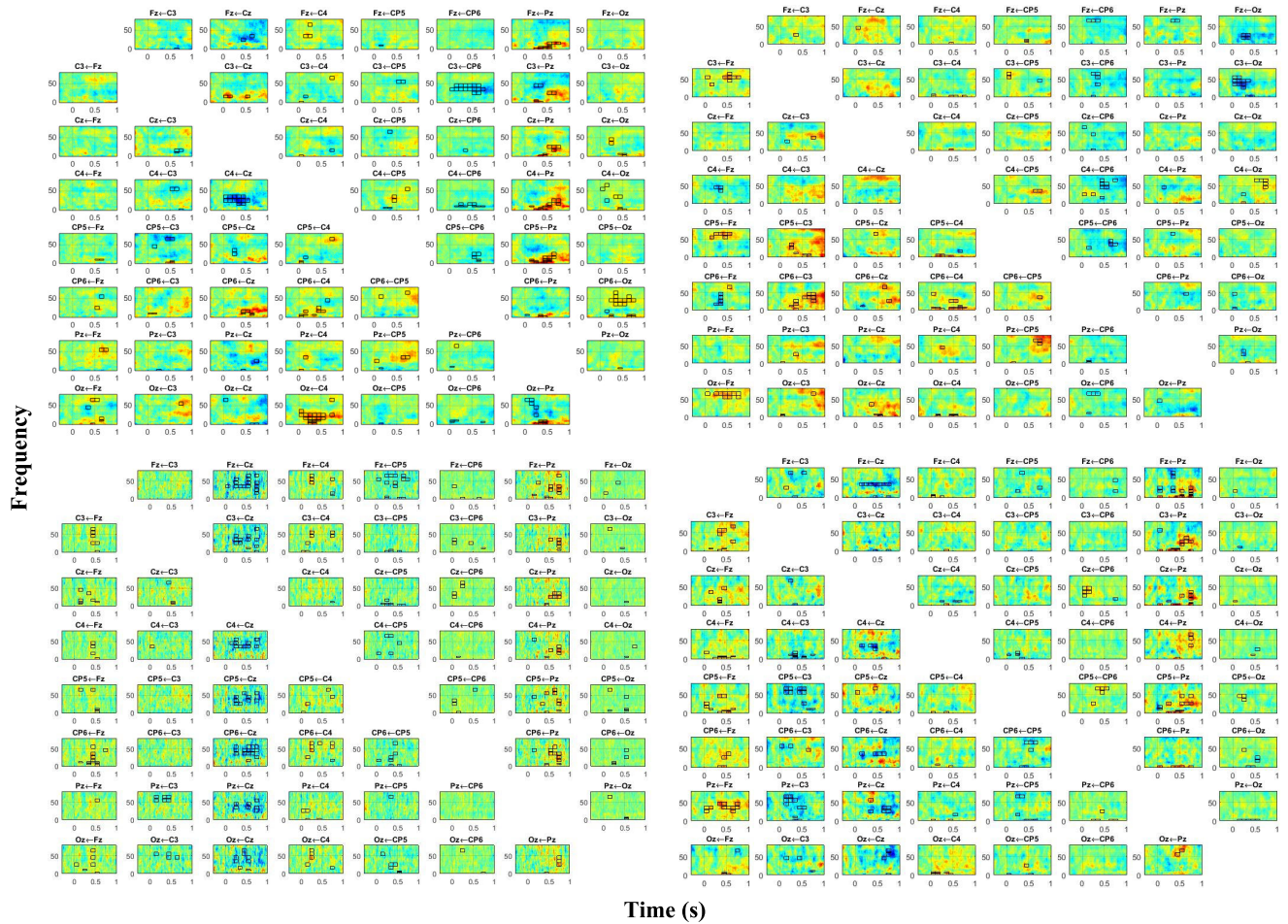
Frequency

Time (s)

FIGURE 3. Differences between the information flows in the target and standard conditions within a time duration of −300 to 1000 ms post-stimulus onset and a frequency range of 0–80 Hz. Red and blue areas denote the information transfers in target condition, which are higher and lower than those in the standard condition, respectively. Black boxes indicate significant difference which have p-values smaller than 0.05.

From the 7th column of each figure, it can be seen that both the STDFT and AF-KF-VNM algorithms yield significant in- and out-going effective information flows at Pz. However, in contrast, no significant connections related to Pz can be captured by the AAR method (see the 7th column in Fig. 3b). Since it has been confirmed that the parietal lobule (Pz) plays a crucial role in P300 generation [54], it is expected that there should be more connections at Pz than the other electrodes. Thus, the results of STDFT and AF-KF-VNM are likely to be more convincing than that obtained by AAR. Moreover, by comparing the results in Fig. 3c and Fig. 3d, it can be seen that the STDTF approach is quite sensitive to the window length selection. It seems that when the window length becomes smaller, the significant connections will spread to wider frequency bands and hence the results will become more dispersive. This may be caused by the larger variance of STDTF as shown in the simulation results in Section IV-A. In contrast, the results of the proposed method can offer more stable results.

Next, we shall focus on the results of the proposed AF-KF-VNM algorithm and study the effective information

flows in separated frequency bands over different intervals. More specifically, the information inside the following frequency bands: δ (0.5-3.5 Hz), θ (4-7 Hz), α (8-13 Hz), $\beta 1$ (14-20 Hz) and $\beta 2$ (21-30 Hz) are integrated (summed) at seven time-intervals with 100ms duration during the whole time course from 0 to 700ms. The choice of the interval length is based on previous studies on the generation of P300 [57], [58]. The results are demonstrated in Fig. 4. It can be seen that most significant connectivity differences occur after 200ms post-stimuli. After that, from 200ms to 800ms, most brain regions are highly controlled by the parietal area (Pz). It can be also noticed that there exists a constant connection from the central area to the occipital area starting from stimuli onset to 800ms. However, in the same duration, the effective information transfers from Cz to C4 and from CP6 to C3 decrease dramatically. Our main observations are summarized in Table 2, with $p < 0.05$.

EEG Source Analysis:

To further validate the above findings, a source analysis is also is performed using the standardized low resolution electromagnetic tomography method (sLORETA) [59] in the

TABLE 2. Main observations and interpretations for the odd ball experiments From EEG signals.

Information Flow	Target Vs Standard Conditions	Duration	Frequency Bands	Spectrum changes in same period	Interpretations/Evidences
C4→others	Increasing	0-100	δ , θ	-	CA control [58], Low frequency modulation by ACC [48].
C4→Oz	Increasing	0-600	$\beta 1$, $\beta 2$, γ	-	Continued movement preparation [61]
C4→CP6	Increasing	0-500	$\beta 1$, $\beta 2$	-	
Cz→C4	Decreasing	0-600	$\beta 2$, γ	Decrease	Inhibitory control [61]
Cz→C3	Increasing	0-200/500-600	$\beta 1$	-	Motor planning and excution [61-63].
Pz→others	Increasing	200-500	δ , θ	Increase	P300 and early ERP components [52]; inhibitory control [64].
Pz→others	Increasing	400-700	α , $\beta 1$, $\beta 2$	Decrease	Late ERP responses [65].
CP6→C3	Decreasing	0-800	γ	Decrease	Reorienting attention [66]; analyzing external signals [67]
CP6→C4	Decreasing	100-800	α , $\beta 1$	-	

Note: γ band denotes frequency larger than 30 Hz. CA: central area. ACC: anterior cingulate cortex.
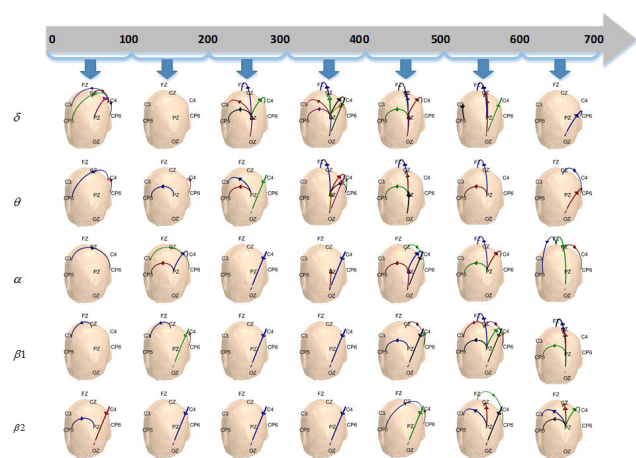

FIGURE 4. Information flows in the target condition, which are significantly stronger (*p*-value<0.05) than those in the standard condition: Arrows indicate the direction of the connections; different connections are denoted with different colors.

Brainstorm toolbox [60]. The source space is generated by SPM with a default 5124-cortex nodes, and the orientation of each source is restricted to be perpendicular to the cortical surface. A three-layer forward model is used to estimate the source signal. The paired sample t-test is performed to compare two conditions across six time durations. The significant difference of the absolute mean value is shown in Fig 2c ($p < 0.05$, FDR corrected).

The target condition always shows stronger activities than the standard condition. More specifically, in the first 100ms, the occipital area (OA) and the central area (CA) have stronger activations in the target condition than in the standard condition, and the difference on CA lasts about 500ms after stimulus onset. This result agrees with the fMRI activations (increased activations at right lingual, left precentral and left postcentral, left supplementary motor areas at target condition) and our connectivity findings, which show that in the first 500ms, the central area always has stronger connections with other areas in target condition. Moreover, the parietal area also shows stronger activities in the target condition

between 100ms to 600ms, which also agrees with our connectivity results and the fMRI results (increased activations at left and right supramargial gyri, left and right angular gyri, right precuneus, left and right inferior parietal gyri, left postcentral area at target condition[2]). It can also be seen that both the source localization and our connectivity analysis better reveal the strong activations at precuneus during 200-600ms, probably due to continual blood supply to this area in both target and non-target conditions, which leads to a reduced difference in activation. On the other hand, both the fMRI and connectivity results demonstrate a significant difference in the left post- and pre- central areas. Moreover, our connectivity analysis shows that the time duration is between 300ms to 600ms, which also agrees with the source localization analysis. However, at time 600ms to 700ms, the excitations for the standard and target stimuli from the source localization analysis are nearly identical and hence there is little difference in activation as shown in the right subplot at the second row of figure 2(b). On the other hand, the connectivity analysis shows that there are still considerable differences in the connectivity between the target and standard conditions. This suggests that the various analyses can complement each other to leverage their strengths and somehow mitigate individual weaknesses, especially on the inherent variations of the data.

C. DISCUSSION AND MAJOR FINDINGS
1) DISCUSSION ON ADAPTIVE FADING KFs

There are several previous works on using adaptive fading to update the covariance matrices [34]–[36]. These adaptive fading KFs only consider the case when there are uncertainties in the state equation of the system model. In practical applications, however, unexpected uncertainties can also exist in the measurement equation. Therefore, their performances may be significantly degraded when the measurement equation experienced considerable uncertainties. To solve this problem,

---

[2]Figure 3a shows the activation on the brain surface. It can be seen that there are also significant activations at left and right supramarginal gyri, angular gyri and left postcentral area. There is also certain activation at right precuneus area.

a more comprehensive approach for systems with inaccurate state or measurement equation was proposed in [37]–[40]. It uses the ratio between the calculated innovation covariance and the estimated one as the fading factor for adaptive compensation of the uncertainties. In this paper, we further employs multiple measurements for further improving its performance [37]–[40]. More simulation results are presented in the supplementary materials. The above experimental results also demonstrated the effectiveness of the proposed AF-KF-VNM algorithm.

2) MAJOR FINDINGS FROM BRAIN SPECTRAL/CONNECTIVITY ANALYSIS

The P300 neuro responses can be divided into three stages: i) information integration (0-200ms), ii) decision process (200ms to P300 latency) and iii) neuronal response (after P300 peak latency) [68]. From Figs. 3 and 4, it can be seen that brain-connection differences located in the information integration stage are fewer than those in other stages. This suggests that, during the information integration stage, most of the brain connectivities in the target condition are similar to that in the standard. This observation is consistent with the previous study in [58]. It can also be noticed that the ERP response differences (at around 300ms) are mainly generated from two brain areas: the parietal area and central area. Next, we shall elaborate our findings separately for these two regions.

a: CENTRAL AREA CONTROL IN EARLY ERP

The fMRI activation map in Fig. 2a shows that the central area (CA) was more active during the target condition, including left pre- and post- central gyri, and right-precentral gyrus. Both our connectivity and source analysis results also show a stronger activation around the left CA in target conditions at 200 to 600ms, probably due to the pressing of the button at the target condition and all the participants are right handed. Li *et al.* [58] also found that the central area played a crucial role in the low frequency components in the ERP responses and suggested that it may control other brain regions by sending out related activation commands in the early ERP stage.

Another EEG-fMRI study [48] of the visual and auditory oddball paradigms reported that there is a significant correlation between the EEG and fMRI signals at left/right anterior cingulate cortex (ACC), left caudate, left precentral gyrus and left middle temporal gyrus around 200ms post-stimulus in the target condition, compared with the standard condition. Since the positions of our C3 and C4 electrodes are near the precentral area and they are also activated during the target task during the same time interval, it suggests that the electrical activations around these regions are likely to be captured by these electrodes. In our EEG connectivity analysis, we find a significantly increase in outgoing-flow from C4 to electrodes at other brain areas, including CP5, CP6 and Pz at parietal and C3 near left precentral areas, in the early 100 milliseconds (the first graph in Fig. 3). Moreover,

prominent information transfers from C4 to other electrodes are found in the delta band during the first 100ms as shown in the first plot of Fig. 4.

Our connectivity analysis also reveals stronger in- and outgoing flows from C4 to other electrodes in the target condition than that in the standard condition. More precisely, there is also a continuous connection from C4 to Oz in both the alpha and beta bands after the onset of the stimulus (approximately from 200ms to 600ms). There is also considerable inflow from Pz to C4 during the about interval. Moreover, in [48], there is also correlates between the EEG with fMRI signals at 425ms and 475ms in right inferior frontal gyrus, middle frontal gyrus, pre- and post- central areas. Also, ACC and anterior supramarginal gyrus respectively correlate with the EEG respectively at 425ms and 475ms together with some other areas at temporal and frontal poles. Our connectivity analysis around 400-500 ms also reveal substantial interactions of the right hemisphere, which aligns with the observations in [48].

From the power spectrum analysis of C4 and Cz in Fig. 2(b), we also find that the target response has significant larger power than standard response at around 4Hz, (i.e. delta band) in the first 100 ms (stage 1), and also significant higher and lower power respectively in stages 2 and 3 at lower frequency (theta, delta and alpha) bands and higher frequency (beta and gamma) bands.

In summary, the fMRI analysis suggests that there is considerable activity difference during the 200ms post-stimulus interval in target condition and the EEG signals recorded by a nearby electrode C4 also show significant differences in power between target and standard responses at around 4Hz, (i.e. delta band) in the first 100 ms (stage 1), and also significant higher and lower power in stages 2 and 3 at theta, delta and alpha bands and beta and gamma bands respectively. Therefore, the two analyses are highly consistent with each other and the EEG spectrum and connectivity further reveal the frequency bands and their time variations associated with the stated activation. The results also indicate that the difference between the target and standard stimuli may be recognized in a very short duration (less than 100 milliseconds) after stimulus onset.

Moreover, as noticed in row 4 from Fig. 4, a significant connection from Cz to C3 arises in the $\beta 1$ frequency band at the first 200 ms as well as in the interval between 500–600 ms, while the connection from Cz to C4 is weakened substantially as shown in Fig. 3(a) (3rd row, 4th column). This is in agreement with the fact that the movement of the right hand is controlled by the left cerebral hemisphere of the brain, which also leads to a decreased activation at C4 in the right hemisphere.

In fact, relevant studies [62], [63] had shown that an event-related desynchronization (ERD) at electrode C3 on the left-brain was observed during the movement of the right-hand. Such activation of the left central area is also found in our fMRI and analysis. This is because subjects are required to press a button using their right index fingers in the target

condition. Moreover, our connectivity analysis suggests that it may be induced by the information flow coming from Cz at the $\beta 1$ frequency band during the 400 to 600ms interval. Though it is expected that there will be a similar ERD at C3 in our spectral analysis, no significant decreased in alpha band power (ERD) is found at C3. Instead, slight ERD in both beta and alpha are observed at C5. This is probably due to subject variability and limited sensitivity of our spectral analysis.

b: PARIETAL AREA CONTROL IN P300 AND LATE ERP

In the fMRI activation analysis, no significant difference was found between the target and standard conditions in the middle-parietal area (though there are activation at right precuneus and other inferior parietal areas), though EEG signals show a strong positive wave (P300) around Pz as shown in Fig. 2(b).

The information flows outgoing from Pz as seen in the seventh column in Fig. 3a are much higher than those from other electrodes in the decision-process and neuronal-response stages. These significantly increased connectivities occur at around 200ms after the onset of the stimuli and last to about 700ms. Similar observations were also demonstrated in [68]. Moreover, from plots in the 7th column of Fig. 3a, we can also find that the information transfer from Pz to other electrodes during the period from 200 to 500 ms post-stimulus are mainly located in the low-frequency band (below 12 Hz). Furthermore, from 400ms to 700ms, the frequency range of connections from Pz to other electrodes is increased to a higher frequency band centered at around 20 Hz. The power spectrum at Pz in Fig. 2(b) also shows significant differences between the target and standard conditions: significant higher low-frequency power and lower alpha-beta bands power in the target condition during the ERP responses.

Two similar components, namely, the early ERP component (with increased low-frequency power) and the ERD or alpha/beta-blocking component (with decreased alpha-beta bands power) were also observed in the analysis of the ERP spectrum [32], [33]. However, the associated connectivity information has not been revealed before. Even though these two components had opposite changes in the spectrum analysis, we found that the outgoing flows from Pz were both increased. The stronger information flows from Pz to other electrodes in the target condition may indicate the connection changes between the parietal and other brain areas. One of the overarching hypothesis of P300 mechanisms is the neuro-inhibition hypothesis, which suggests that the P300 could represent rapid neural inhibition of ongoing activity [64]. Coupled with our observations, such inhibition mechanism may be mostly controlled by the parietal lobule.

All these findings suggest that the proposed AF-KF-VNM-based spectrum and functional connectivity analyses of EEG can complement the fMRI analysis and provide more information on the detailed frequency components, their time variations and possible information flows between different parts of the brain. It thus serves as a promising alternative for connectivity analysis of nonstationary

multi-channel EEGs and other related signals. Moreover, the AF-KF-VNM algorithm is very general and may find other applications in the analysis of biomedical signals.

Previous studies have reported the effect of possible volume condition (VC) effects in the estimation of DTF from EEG electrodes [69], [70]. While it is truth that the DTF results based on EEG scalp data is not completely free from this effect, DTF based on MVAR methods are less insensitive to VC comparing to classic connectivity methods and the VC effect does not ruin critically DTF results [71]. Nunez and Srinivasan [72] studied both theoretically and experimentally the VC effect using coherence technique. The results showed that the VC effect were close to zero when the distance between two electrodes reached about 7cm. Thus, for the 10-20 system we used in this study, the effect of VC on the connectivity results is likely to be limited.

V. CONCLUSION

A novel AF-KF-VNF-based approach to TV brain spectral and functional connectivity analyses of event-related multi-channel EEG signals has been presented. It models the EEG signals as a TV MVAR process for estimating the spectrum of the EEG signals and identifying their functional connectivity. The proposed AF-KF-VNM algorithm uses a new AF method to adaptively update the model parameters of the KF and utilizes multiple measurements for respectively automatic compensation of possibly imperfect model parameters in the KF and better adaptation to nonstationary signals. Experimental results on a simulated data for modeling the TV directed interactions in neural data show that the proposed method yield better tracking performance than other approaches tested. A novel methodology for combined fMRI activation maps and EEG spectrum analyses has also been presented for quantifying the differences in spectrum contents and information flows between the target and standard conditions in a visual oddball paradigm. The results and findings show that the proposed methodology agrees well with the literature and is able to reveal significant frequency components and information flow involved as well as their time variations. It thus serves as a promising alternative for spectral and connectivity analyses of nonstationary biomedical signals like multi-channel EEG.

REFERENCES

[1] M. Case, "Characterization of functional brain activity and connectivity using EEG and fMRI in patients with sickle cell disease," *NeuroImage: Clin.*, vol. 14, pp. 1–17, Dec. 2017, doi: 10.1016/j.nicl.2016.12.024.

[2] S. Mostafa, L. Tang, and F.-X. Wu, "Diagnosis of autism spectrum disorder based on eigenvalues of brain networks," *IEEE Access*, vol. 7, pp. 128474–128486, 2019.

[3] S. Hanouneh, H. U. Amin, N. M. Saad, and A. S. Malik, "EEG power and functional connectivity correlates with semantic long-term memory retrieval," *IEEE Access*, vol. 6, pp. 8695–8703, 2018.

[4] G. Wang, D. Ren, K. Li, D. Wang, M. Wang, and X. Yan, "EEG-based detection of epileptic seizures through the use of a directed transfer function method," *IEEE Access*, vol. 6, pp. 47189–47198, 2018.

[5] T. Zhang, M. Li, L. Zhang, B. Biswal, D. Yao, and P. Xu, "The time-varying network patterns in motor imagery revealed by adaptive directed transfer function analysis for fMRI," *IEEE Access*, vol. 6, pp. 60339–60352, 2018.

[6] E. Pereda, R. Q. Quiroga, and J. Bhattacharya, "Nonlinear multivariate analysis of neurophysiological signals," *Progr. Neurobiol.*, vol. 77, nos. 1–2, pp. 1–37, Sep. 2005, doi: 10.1016/j.pneurobio.2005.10.003.

[7] K. Pearson, "Note on Regression and Inheritance in the Case of Two Parents," *Proc. Roy. Soc. London*, vol. 58, no. 1, pp. 240–242, 2006.

[8] Z. G. Zhang, X. L. Cai, S. C. Chan, Y. Hu, L. Hu, and C. Q. Chang, "Time-frequency coherence analysis of multi-channel eventrelated potential using adaptive windowed Lomb periodogram," in *Proc. 4th Int. IEEE/EMBS Conf. Neural Eng.*, Apr. 2009, pp. 657–660.

[9] R. Srinivasan, W. R. Winter, J. Ding, and P. L. Nunez, "EEG and MEG coherence: Measures of functional connectivity at distinct spatial scales of neocortical dynamics," *J. Neurosci. Methods*, vol. 166, no. 1, pp. 41–52, Oct. 2007.

[10] M. Ding, Y. Chen, and S. L. Bressler, "Granger causality: Basic theory and application to neuroscience," in *Handbook of Time Series Analysis*, S. Schelter, M. Winterhalder, J. Timmer, Eds. Wienheim, Germany: Wiley, 2006, pp. 438–460.

[11] L. A. Baccalá and K. Sameshima, "Partial directed coherence: A new concept in neural structure determination," *Biol. Cybern.*, vol. 84, no. 6, pp. 463–474, May 2001, doi: 10.1007/PL00007990.

[12] M. Kamiński, M. Ding, W. A. Truccolo, and S. L. Bressler, "Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance," *Biol. Cybern.*, vol. 85, no. 2, pp. 145–157, Aug. 2001, doi: 10.1007/s004220000235.

[13] M. J. Kaminski and K. J. Blinowska, "A new method of the description of the information flow in the brain structures," *Biol. Cybern.*, vol. 65, no. 3, pp. 203–210, Jul. 1991.

[14] T. Amemiya, "Multivariate regression and simultaneous equation models when the dependent variables are truncated normal," *Econometrica*, vol. 42, no. 6, pp. 999–1012, Nov. 1974.

[15] F. Babiloni, F. Cincotti, C. Babiloni, F. Carducci, D. Mattia, L. Astolfi, A. Basilisco, P. M. Rossini, L. Ding, Y. Ni, J. Cheng, K. Christine, J. Sweeney, and B. He, "Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function," *NeuroImage*, vol. 24, no. 1, pp. 118–131, Jan. 2005.

[16] A. Sohrabpour, S. Ye, G. A. Worrell, W. Zhang, and B. He, "Noninvasive electromagnetic source imaging and granger causality analysis: An electrophysiological connectome (eConnectome) approach," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 12, pp. 2474–2487, Dec. 2016.

[17] J. Ginter, M. Kamiński, P. J. Durka, G. Pfurtscheller, C. Neuper, and K. J. Blinowska, "Propagation of EEG activity in the beta and gamma band during movement imagery in humans," *Methods Inf. Med.*, vol. 44, no. 1, pp. 106–113, 2005.

[18] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, Mar. 1960.

[19] M. Arnold, X. H. R. Milner, H. Witte, R. Bauer, and C. Braun, "Adaptive AR modeling of nonstationary time series by means of Kalman filtering," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 5, pp. 553–562, May 1998.

[20] M. P. Tarvainen, J. K. Hiltunen, P. O. Ranta-Aho, and P. A. Karjalainen, "Estimation of nonstationary EEG with Kalman smoother approach: An application to event-related synchronization (ERS)," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 3, pp. 516–524, Mar. 2004, doi: 10.1109/TBME.2003.821029.

[21] A. Schlögl, D. Flotzinger, and G. Pfurtscheller, "Adaptive autoregressive modeling used for single-trial EEG classification," *Biomed. Tech. Biomed. Eng.*, vol. 42, no. 6, pp. 162–167, 1997.

[22] M. J. Cassidy and W. D. Penny, "Bayesian nonstationary autoregressive models for biomedical signal analysis," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 10, pp. 1142–1152, Oct. 2002, doi: 10.1109/TBME.2002.803511.

[23] S. M. Kay, *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1988.

[24] T. Milde, L. Leistritz, L. Astolfi, W. H. R. Miltner, T. Weiss, F. Babiloni, and H. Witte, "A new Kalman filter approach for the estimation of high-dimensional time-variant multivariate AR models and its application in analysis of laser-evoked brain potentials," *NeuroImage*, vol. 50, no. 3, pp. 960–969, Apr. 2010, doi: 10.1016/j.neuroimage.2009.12.110.

[25] J. Toppi, F. Babiloni, G. Vecchiato, F. De Vico Fallani, D. Mattia, S. Salinari, T. Milde, L. Leistritz, H. Witte, and L. Astolfi, "Towards the time varying estimation of complex brain connectivity networks by means of a General Linear Kalman Filter approach," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug./Sep. 2012, pp. 6192–6195, doi: 10.1109/EMBC.2012.6347408.

[26] M. F. Pagnotta and G. Plomp, "Time-varying MVAR algorithms for directed connectivity analysis: Critical comparison in simulations and benchmark EEG data," *PLoS ONE*, vol. 13, no. 6, Jun. 2018, Art. no. e0198846, doi: 10.1371/journal.pone.0198846.

[27] A. H. Omidvarnia, M. Mesbah, M. S. Khlif, J. M. O'Toole, P. B. Colditz, and B. Boashash, "Kalman filter-based time-varying cortical connectivity analysis of newborn EEG," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug./Sep. 2011, pp. 1423–1426, doi: 10.1109/IEMBS.2011.6090335.

[28] K. J. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *NeuroImage*, vol. 19, no. 4, pp. 1273–1302, Aug. 2003, doi: 10.1016/s1053-8119(03)00202-7.

[29] O. David, S. J. Kiebel, L. M. Harrison, J. Mattout, J. M. Kilner, and K. J. Friston, "Dynamic causal modeling of evoked responses in EEG and MEG," *NeuroImage*, vol. 30, no. 4, pp. 1255–1272, May 2006, doi: 10.1016/j.neuroimage.2005.10.045.

[30] A. Schlögl, *The Electroencephalogram and the Adaptive Autoregressive Model: Theory and Applications*. Aachen, Germany: Shaker Verlag, 2000.

[31] M. Niedzwiecki, *Identification of Time-Varying Processes*. Hoboken, NJ, USA: Wiley, 2000.

[32] Z. G. Zhang, S. C. Chan, and X. Chen, "A new Kalman filter-based recursive method for measuring and tracking time-varying spectrum of nonstationary signals," in *Proc. 9th Int. Conf. Inf., Commun. Signal Process.*, Dec. 2013, pp. 1–4.

[33] Z. G. Zhang, K. M. Tsui, S. C. Chan, W. Y. Lau, and M. Aboy, "A novel method for nonstationary power spectral density estimation of cardiovascular pressure signals based on a Kalman filter with variable number of measurements," *Med. Biol. Eng. Comput.*, vol. 46, no. 8, pp. 789–797, Aug. 2008.

[34] Q. Xia, M. Rao, Y. Ying, and X. Shen, "Adaptive fading Kalman filter with an application," *Automatica*, vol. 30, no. 8, pp. 1333–1338, Aug. 1994.

[35] C. Hide, T. Moore, and M. Smith, "Adaptive Kalman filtering for low-cost INS/GPS," *J. Navigat.*, vol. 56, no. 1, pp. 143–152, Jan. 2003.

[36] M. Efe and L. Ozbek, "Fading Kalman filter for manoeuvring target tracking," *Istatistik J. Turkish Stat. Assoc.*, vol. 2, no. 3, pp. 193–206, 2012.

[37] K.-H. Kim, G.-I. Jee, and J.-H. Song, "Carrier tracking loop using the adaptive two-stage Kalman filter for high dynamic situations," *Int. J. Control, Autom., Syst.*, vol. 6, no. 6, pp. 948–953, 2008.

[38] K.-H. Kim, G.-I. Jee, C.-G. Park, and J.-G. Lee, "The stability analysis of the adaptive fading extended Kalman filter using the innovation covariance," *Int. J. Control, Autom. Syst.*, vol. 7, no. 1, pp. 49–56, Feb. 2009.

[39] K. Hoon Kim, J. Gyu Lee, and C. Gook Park, "Adaptive two-stage extended Kalman filter for a fault-tolerant INS-GPS loosely coupled system," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, no. 1, pp. 125–137, Jan. 2009.

[40] K. H. Kim, J. G. Lee, and C. G. Park, "Adaptive two-stage Kalman filter in the presence of unknown random bias," *Int. J. Adapt. Control Signal Process.*, vol. 20, no. 7, pp. 305–319, Sep. 2006.

[41] C.-M. Ting, S.-H. Salleh, Z. Z. Zainuddin, and A. Bahar, "Spectral estimation of nonstationary EEG using particle filtering with application to event-related desynchronization (ERD)," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 2, pp. 321–331, Feb. 2011.

[42] L. Sommerlade, K. Henschel, J. Wohlmuth, M. Jachan, F. Amtage, B. Hellwig, C. H. Lücking, J. Timmer, and B. Schelter, "Time-variant estimation of directed influences during parkinsonian tremor," *J. Physiol.-Paris*, vol. 103, no. 6, pp. 348–352, Nov. 2009, doi: 10.1016/j.jphysparis.2009.07.005.

[43] *Openfmri Database*. Accessed: Sep. 2017. [Online]. Available: https://openfmri.org/dataset/ds000116/

[44] S. C. Kleih, "Out of the frying pan into the fire—The P300-based BCI faces real-world challenges," *Prog. Brain Res.*, vol. 194, no. 1, pp. 27–46, 2011.

[45] Y. Li, J. Pan, F. Wang, and Z. Yu, "A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 11, pp. 3156–3166, Nov. 2013, doi: 10.1109/TBME.2013.2270283.

[46] T. W. Picton, "The P300 wave of the human event-related potential," *J. Clin. Neurophysiol.*, vol. 9, no. 4, pp. 456–479, Oct. 1992.

[47] V. P. Clark, S. Fannon, S. Lai, R. Benson, and L. Bauer, "Responses to rare visual target and distractor stimuli using event-related fMRI," *J. Neurophysiol.*, vol. 83, no. 5, pp. 3133–3139, May 2000.

[48] J. M. Walz, R. I. Goldman, M. Carapezza, J. Muraskin, T. R. Brown, and P. Sajda, "Simultaneous EEG-fMRI reveals temporal evolution of coupling between supramodal cortical attention networks and the brainstem," *J. Neurosci.*, vol. 33, no. 49, pp. 19212–19222, Dec. 2013.

[49] W. Lou, J. Xu, H. Sheng, and S. Zhao, "Multichannel linear descriptors analysis for event-related EEG of vascular dementia patients during visual detection task," *Clin. Neurophysiol.*, vol. 122, no. 11, pp. 2151–2156, Nov. 2011.

[50] D. E. J. Linden, "The p300: Where in the brain is it produced and what does it tell us?" *Neuroscientist*, vol. 11, no. 6, pp. 563–576, Dec. 2005.

[51] C. Bledowski, "Localizing P300 generators in visual target and distractor processing: A combined event-related potential and functional magnetic resonance imaging study," *J. Neurosci.*, vol. 24, no. 42, pp. 9353–9360, Oct. 2004, doi: 10.1523/JNEUROSCI.1897-04.2004.

[52] Z. G. Zhang, Y. S. Hung, and S. C. Chan, "Local polynomial modeling of time-varying autoregressive models with application to Time–Frequency analysis of event-related EEG," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 557–566, Mar. 2011.

[53] J. B. Kadane and N. A. Lazar, "Methods and criteria for model selection," *J. Amer. Statist. Assoc.*, vol. 99, pp. 279–290, Mar. 2004.

[54] G. McCarthy, M. Luby, J. Gore, and P. Goldman-Rakic, "Infrequent events transiently activate human prefrontal and parietal cortex as measured by functional MRI," *J. Neurophysiol.*, vol. 77, no. 3, pp. 1630–1634, Mar. 1997.

[55] V. Menon, J. M. Ford, K. O. Lim, G. H. Glover, and A. Pfefferbaum, "Combined event-related fMRI and EEG evidence for temporal—Parietal cortex activation during target detection," *Neuroreport*, vol. 8, no. 14, pp. 3029–3037, 1997.

[56] E. Maris and R. Oostenveld, "Nonparametric statistical testing of EEG- and MEG-data," *J. Neurosci. Methods*, vol. 164, no. 1, pp. 177–190, Aug. 2007, doi: 10.1016/j.jneumeth.2007.03.024.

[57] R. Verleger, W. Heide, C. Butt, and D. Kömpf, "Reduction of P3b in patients with temporo-parietal lesions," *Cognit. Brain Res.*, vol. 2, no. 2, pp. 103–116, Sep. 1994.

[58] F. Li, B. Chen, H. Li, T. Zhang, F. Wang, Y. Jiang, P. Li, T. Ma, R. Zhang, Y. Tian, T. Liu, D. Guo, D. Yao, and P. Xu, "The time-varying networks in P300: A task-evoked EEG study," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 7, pp. 725–733, Jul. 2016, doi: 10.1109/TNSRE.2016.2523678.

[59] R. D. Pascual-Marqui, "Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details," *Methods Find Exp. Clin. Pharmacol.*, vol. 24, pp. 5–12, Jan. 2002.

[60] *Brainstorm*. Sep. 2018. [Online]. Available: http://neuroimage.usc.edu/brainstorm/

[61] Y. Zhang, Y. Chen, S. L. Bressler, and M. Ding, "Response preparation and inhibition: The role of the cortical sensorimotor beta rhythm," *Neuroscience*, vol. 156, no. 1, pp. 238–246, Sep. 2008.

[62] V. Morash, O. Bai, S. Furlani, P. Lin, and M. Hallett, "Classifying EEG signals preceding right hand, left hand, tongue, and right foot movements and motor imageries," *Clin. Neurophysiol.*, vol. 119, no. 11, pp. 2570–2578, Nov. 2008.

[63] C. Toro, G. Deuschl, R. Thatcher, S. Sato, C. Kufta, and M. Hallett, "Event-related desynchronization and movement-related cortical potentials on the ECoG and EEG," *Electroencephalogr. Clin. Neurophysiol./Evoked Potentials Sect.*, vol. 93, no. 5, pp. 380–389, Oct. 1994.

[64] J. Polich, "Updating p300: An integrative theory of P3a and P3b," *Clin. Neurophysiol.*, vol. 118, no. 10, pp. 2128–2148, Oct. 2007.

[65] S. Makeig, A. Delorme, M. Westerfield, T.-P. Jung, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Electroencephalographic brain dynamics following manually responded visual targets," *PLoS Biol.*, vol. 2, no. 6, p. e176, 2004.

[66] A. Abu-Akel and S. Shamay-Tsoory, "Neuroanatomical and neurochemical bases of theory of mind," *Neuropsychologia*, vol. 49, no. 11, pp. 2971–2984, Sep. 2011.

[67] S.-J. Blakemore, D. M. Wolpert, and C. D. Frith, "Abnormalities in the awareness of action," *Trends Cognit. Sci.*, vol. 6, no. 6, pp. 237–242, Jun. 2002.

[68] C. Wang, J. Xu, W. Lou, and S. Zhao, "Dynamic information flow analysis in vascular dementia patients during the performance of a visual oddball task," *Neurosci. Lett.*, vol. 580, pp. 108–113, Sep. 2014, doi: 10.1016/j.neulet.2014.07.056.

[69] C. Brunner, M. Billinger, M. Seeber, T. R. Mullen, and S. Makeig, "Volume conduction influences scalp-based connectivity estimates," *Frontiers Comput. Neurosci.*, vol. 10, p. 121, Nov. 2016, doi: 10.3389/fncom.2016.00121.

[70] F. Van de Steen, L. Faes, E. Karahan, J. Songsiri, P. A. Valdes-Sosa, and D. Marinazzo, "Critical comments on EEG sensor space dynamical connectivity analysis," *Brain Topography*, vol. 32, no. 4, pp. 643–654, Jul. 2019, doi: 10.1007/s10548-016-0538-7.

[71] M. Kaminski and K. J. Blinowska, "The influence of volume conduction on DTF estimate and the problem of its mitigation," *Frontiers Comput. Neurosci.*, vol. 11, p. 36, May 2017, doi: 10.3389/fncom.2017.00036.

[72] P. L. Nunez and R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*. New York, NY, USA: Oxford Univ. Press, 2006.

JIEWEI LI received the B.Eng. degree in electrical and electronic engineering from Shandong University, China, in 2010, and the master's degree in medical sciences from The University of Hong Kong, Hong Kong, in 2014, where he is currently pursuing the Ph.D. degree in biomedical signal processing. His research interests include biomedical signal processing and brain–computer interface.

SHING-CHOW CHAN (Member, IEEE) received the B.Sc. (Eng.) and Ph.D. degrees in electrical engineering from The University of Hong Kong, in 1986 and 1992, respectively. Since 1994, he has been with the Department of Electrical and Electronic Engineering, The University of Hong Kong, where he is currently a Professor and the Associate Head of the Department. His research interests include fast transform algorithms, filter design and realization, multirate and biomedical signal processing, communications and array signal processing, high-speed A/D converter architecture, bioinformatics, smart grid, and image-based rendering. He is currently a member of the Digital Signal Processing Technical Committee of the IEEE Circuits and Systems Society. He was the Chair of the IEEE Hong Kong Chapter of Signal Processing from 2000 to 2002, and an Organizing Committee Member of the 2003 IEEE ICASSP and the 2010 IEEE ICIP. He served as an Associate Editor for *Digital Signal Processing*, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I, from 2008 to 2009, and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, from 2012 to 2016. He is an Associate Editor of *Journal of Signal Processing Systems*.

ZHONG LIU received the B.Eng. degree from the Huaihai Institute of Technology, Lianyungang, China, in 2010, the M.Eng. degree from the Zhongyuan University of Technology, Zhengzhou, China, in 2013, and the Ph.D. degree from the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, in 2018. He is currently a Research Fellow at the Department of Biomedical Engineering, Shenzhen University, Shenzhen, China. His research interests include signal processing, pattern recognition, and human–computer interaction.

CHUNQI CHANG (Member, IEEE) received the B.Eng. and M.Eng. degrees in electronic engineering from the University of Science and Technology of China, Hefei, China, in 1992 and 1995, respectively, and the Ph.D. degree in biomedical engineering from The University of Hong Kong, Hong Kong, in 2001. From 2002 to 2012, he was with The University of Hong Kong, as a Postdoctoral Fellow, a Research Associate, a Research Assistant Professor with the Department of Electrical and Electronic Engineering, and a Visiting Professor with the State Key Laboratory of Brain and Cognitive Sciences. From 2012 to 2015, he was a Professor with the School of Electronics and Information Engineering, Soochow University, Suzhou, China. Since 2015, he has been a Professor with the Health Science Center, School of Biomedical Engineering, Shenzhen University, Shenzhen, China. His current research interests include brain decoding, functional brain imaging, neural signal processing, statistical signal processing, and machine learning.

● ● ●