

Adaptive Fourier Decomposition for Multi-Channel Signal Analysis

Ze Wang , Student Member, IEEE, Chi Man Wong , Agostinho Rosa , Tao Qian ,
and Feng Wan , Senior Member, IEEE

Abstract—Evolved from the conventional Fourier decomposition based on a pre-defined basis, Adaptive Fourier decomposition (AFD) uses adaptive basis to achieve the fast energy convergence. This paper extends the AFD to the multi-channel case, which finds common adaptive basis across all channels. The proposed multi-channel AFD (MAFD) scheme includes the multi-channel core AFD for general signals and the multi-channel unwinding AFD for specific signals that have common inner functions. Owing to the merits of the original AFD, the MAFD can provide sparse joint time-frequency distribution by computing the transient time frequency distribution (TTFD) across channels. Simulations on synthetic and real-world signals demonstrate that the proposed scheme can find and apply the common adaptive basis with desired properties maintained by the AFD, showing high potentials in real-world applications.

Index Terms—Amplitude and frequency modulated signal, adaptive Fourier decomposition, multi-channel signal, time-frequency analysis.

I. INTRODUCTION

THE observations in the physical sciences and engineering often form time-varying signals, which cannot be characterized adequately by conventional Fourier analysis [1]. The univariate modulated oscillation model that describes the one dimensional (1D) time-varying signals as amplitude and frequency modulated oscillations provides an attractive representation and has become a standard model. To characterize time-varying signals with arbitrary number of channels, the univariate modulated

oscillation model is extended to the modulated multivariate oscillation model in [2]. This model assumes a common oscillation spanning all of individual channel signals. An important task towards multi-channel signal analysis is to identify the common oscillatory structure [2], [3].

As the interest in the multi-channel signal analysis grows, multi-channel extensions of several conventional algorithms have been proposed to explore dependencies among multiple channels. The work in [2] extended the wavelet ridge transform for the modulated multivariate oscillations in which the multivariate wavelet transform (MWT) identifies the local maximal points for computing the scale parameter in the wavelet coefficients and thus can provide local oscillatory dynamics of multi-channel signals. To reduce the restrictions due to the linear projection in the wavelet transform, the synchrosqueezing transform is introduced to the MWT in [3]. The multivariate synchrosqueezing wavelet transform (MSWT) reassigns the energies of decomposition components to make them concentrated around the instantaneous frequency curves of the modulated oscillations. Unfortunately, the wavelet transform is based on the pre-defined mother wavelet. In order to decompose signals adaptively, the empirical mode decomposition (EMD) is proposed and later extended as the multivariate EMD (MEMD) for analyzing multi-channel signals [4]. Instead of using the EMD to analyze signals separately on each channel, the MEMD analyzes signals in all channels simultaneously and can produce same number of intrinsic mode functions (IMFs) adaptively for all channels. To enhance the computational efficiency, the fast MEMD (FMEMD) was proposed [5]. Although the MEMD performs well on the analyses of the synthetic sinusoidal signals and real-world signals [6]–[8], it lacks mathematical theory and necessary theoretical guarantees. To tackle this issue, several novel adaptive decomposition methods were proposed in literature. Based on the Fourier theory and zero-phase filtering as well as the discrete cosine and sine transforms, the Fourier decomposition method (FDM), the Fourier quadrature transform (FQT) and the corresponding multi-channel extension called multivariate FDM (MFDM), were introduced [9], [10]. In the MFDM, the frequency-domain components are rearranged by zero-phase filter banks to form band limited Fourier intrinsic band functions (FIBFs) for analyzing multivariate nonlinear and non-stationary time series. Except the MFDM, the multivariate variational mode decomposition (MVMD), originated from the variation mode decomposition (VMD), achieves the decompositions of the multivariate data by minimizing the collective bandwidth of modes whose frequency bands limits around the center frequency [11], [12]. Then, to better process wide-band and time-varying multi-channel signals, the multivariate versions of the nonlinear chirp mode

Manuscript received May 10, 2021; revised November 16, 2021; accepted January 7, 2022. Date of publication January 18, 2022; date of current version February 23, 2022. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Athanasios A. Rontogiannis. This work was supported in part by the Science and Technology Development Fund, Macau SAR under Grants 055/2015/A2, 0045/2019/AFJ, 0018/2019/AKP, 0123/2018/A3 and 0022/2021/APD, in part by the University of Macau under Grants MYRG2016-00240-FST and MYRG2017-00207-FST, and in part by the LARSyS - FCT Project under Grant UIDB/50009/2020. (Corresponding author: Feng Wan.)

Ze Wang, Chi Man Wong, and Feng Wan are with the Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Taipa, Macau 999078, China, with the Centre for Cognitive and Brain Sciences, and also with the Centre for Artificial Intelligence and Robotics, Institute of Collaborative Innovation, University of Macau, Taipa, Macau (e-mail: wangze19910407@gmail.com; chiman465@gmail.com; fwan@um.edu.mo).

Agostinho Rosa is with the Department of Bioengineering, LaSEEB-System and Robotics Institute, Instituto Superior Tecnico, University of Lisbon, Lisbon 1049-001, Portugal (e-mail: acrosa@laseeb.org).

Tao Qian is with the Macao Center for Mathematical Sciences, Macau University of Science and Technology, Taipa, Macau (e-mail: tqian@must.edu.mo).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TSP.2022.3143723>, provided by the authors.

Digital Object Identifier 10.1109/TSP.2022.3143723

decomposition (NCMD), i.e., the multivariate nonlinear chirp mode decomposition (MNCMD) and the multivariate intrinsic chirp mode decomposition (MICMD), can transform the multivariate nonlinear chirp modes into narrow-band multivariate signals by minimizing the sum of bandwidths of modes across channels [13], [14]. Compared with the MEMD, these novel adaptive multi-channel decomposition methods include more mathematical theorems or mathematical analysis to avoid the limitations of the EMD with regard to the mode mixing, the detrend uncertainty and the end effect artifacts [9], [12], [14], [15]. However, compared with the basis based decomposition methods, such as the conventional Fourier decomposition, the decomposition components of these novel adaptive multi-channel decomposition methods still lack the support of the rigorous mathematical foundation and cannot be formulated by mathematical expressions. This issue makes characteristics of the decomposition components difficult to be interpreted and analyzed mathematically, which limits their applications in practice. In this study, the MEMD, the MFDM, the MVMD and the MNCMD are collectively called empirical decomposition methods. The AFD, combining the advantages of the basis based and empirical decomposition methods, was proposed [16].

The AFD provides a greedy iterative decomposition of time-varying signals into a series of mono-components (MCs) only containing non-negative analytic phase derivatives [16]. The decomposition of the AFD is still based on basis. However, unlike the conventional basis based decomposition methods using the pre-defined basis, the AFD uses adaptive basis to decompose signals so that it can achieve the fast energy convergence [15]. Until now, for 1D signal, there are several algorithms implementing the AFD. The core AFD is the most fundamental AFD method. It applies the modified Blaschke product as its basis and adopts the matching pursuit decomposition process. In each decomposition level, one basis component is searched from the dictionary to extract the largest possible energy portion from the remainder. Once a basis component is obtained, it will not be adjusted in the subsequent decomposition processes. Based on such iterative decomposition procedures, the core AFD can achieve the largest energy shrinkage in every decomposition level [16]. However, it cannot guarantee the global optimal convergence efficiency. To improve the global convergence efficiency of the core AFD, the cyclic AFD is proposed. Assuming that the total number of basis components is known, then the basis components obtained from the core AFD can be adjusted one by one to maximize the total extracted energy [17]. The cyclic AFD adjusts one basis component by fixing other basis components and applying the same basis searching scheme as in the core AFD. This adjustment process is repeated for all basis components until the total energy of all extracted components achieves its maximum value. Besides the cyclic AFD, the unwinding AFD is also proposed for speeding up the convergence. Although the unwinding AFD also adopts the matching pursuit process, it combines the factorization process [18]. By factorizing the inner function factor from the remainder, the energy convergence can practically increase the convergence rate. Using the core AFD as foundation, these AFD methods adopt different basis searching processes under different assumptions to achieve the fast energy convergence. Moreover, because all decomposed components are MCs, the transient time-frequency distribution (TTFD) generated based on the AFD satisfies many good mathematical properties, including correct total energy, non-negative,

real-valuedness, weak and strong finite supports [15]. The AFD is not only proposed for the 1D signals, but also extended to the two dimensional (2D) and higher dimensional signals [19]. The multi-dimensional AFD applies the Kronecker product of multiple finite Takenaka-Malmquist (TM) systems in multiple dimensions as its basis. Based on the similar decomposition procedure of the 1D AFD, the multi-dimensional AFD can express the processed multi-dimensional signal to the combination of multi-dimensional MCs with the high efficiency of the energy shrinkage. The AFD has shown good performance in system identification, modeling, signal compression and denoising [20]–[25].

In order to apply the main characteristics of the AFD for analyzing the multi-channel signals, the extension of the AFD to the multi-channel version is a natural and desired development. Unfortunately, the existing AFD methods only work for single channel signals. It is well-known that multi-channel data usually shares common oscillations across channels. If simply applying the AFD separately on each channel, the common knowledge between channels will be neglected, which may yield different numbers of misaligned MCs for different channels and further lead to physically meaningless estimates or decisions. On the other hand, multi-channel signals can be considered as multi-dimensional signals and thus analyzed by the multi-dimensional AFD. The existing multi-dimensional AFD assumes that the decomposition components are continuous in all dimensions to search adaptive basis [19], [26]. However, the decomposition of multi-channel signals should not be constrained by the continuity in the channel dimension.

The main contributions of this study are that i) a multi-channel AFD (MAFD) is proposed, which is the first extension of the AFD for analyzing multi-channel signals; ii) fast Fourier transform (FFT) based modifications are applied to improve the computation efficiency of the proposed MAFD; iii) a joint time-frequency distribution is further developed based on the MAFD to provide the coherent time-frequency analysis for the multi-channel signals; iv) the characteristics of the proposed MAFD as well as the good time-frequency localization ability of the proposed joint time-frequency distribution are illustrated through simulations on synthetic and real-world multi-channel signals. The proposed MAFD contains several good characteristics including i) applying the common basis components for all channels to achieve the alignment of common or joint oscillations across channels; ii) introducing characteristics of the AFD to the analysis of multi-channel signals, which are the fast energy convergence with adaptive basis components and the rigorous mathematical foundation as well as the decomposition components with non-negative and continuous phase derivatives; iii) owing to these inherited characteristics from the AFD, being able to provide the joint time-frequency distribution that follows mathematical properties of the TTFD. In multivariate oscillations, the bivariate and trivariate cases are most important [2], [3]. We focus on analyzing the multi-channel 1D signals (bivariate oscillations) in this paper. The similar analysis can be applied to the multi-channel 2D signals (trivariate oscillations).

The organization of the rest part in this paper is as follows: Section II gives the key notations used in this paper, and introduces the modulated multivariate oscillations and the conventional adaptive Fourier decomposition. Section III presents the proposed multi-channel extensions of the core AFD and the unwinding AFD as well as the joint time-frequency analysis based on the MAFD. Section IV verifies the MAFD through

simulations and examples of real-world applications. Then, the basis searched by the proposed MAFD and the computation complexity of the proposed implementation of the MAFD as well as the comparisons between the proposed MAFD and the empirical decomposition methods are discussed in Section V. Finally, the conclusion is drawn in Section VI.

II. PRELIMINARIES

A. Notations

For a clear introduction of the MAFD methods, Table I summarizes the key notations used in this paper. In general, the variables with right subscripts c and n denote that they correspond to the c -th channel and the n -th decomposition level. In addition, the function notations with the parentheses and the square brackets denote the continuous signal and the corresponding discrete signal. For example, $G_{c,n}(e^{jt})$ and $G_{c,n}[k]$ are the continuous and corresponding discrete processed analytic signal of the c -th channel in the n -th decomposition level, and have the relationship $G_{c,n}[k] = G_{c,n}(e^{jt_k})$ with the k -th sampling time t_k .

B. Modulated Multivariate Oscillations

For a signal $x(t)$ containing time-varying amplitudes and frequencies, the corresponding analytic signal $x_+(t)$ can be described by the modulated oscillation model

$$x_+(t) = x(t) + j\mathcal{H}\{x(t)\} = \zeta(t)e^{j\psi(t)} \quad (1)$$

where $\zeta(t)$ and $\psi(t)$ are the instantaneous amplitude and phase respectively, j is the imaginary unit, and $\mathcal{H}(\cdot)$ denotes the Hilbert transform. This concept of univariate modulated oscillation has been extended to the multivariate case, which models the joint oscillatory structure of the multi-channel signal [2]. For a multi-channel signal $\mathbf{x}(t)$, we can extend the representation in (1) to describe the corresponding multi-channel analytic signal by the modulated multivariate oscillation

$$\mathbf{x}_+(t) = \begin{bmatrix} \zeta_1(t)e^{j\psi_1(t)} \\ \zeta_2(t)e^{j\psi_2(t)} \\ \vdots \\ \zeta_C(t)e^{j\psi_C(t)} \end{bmatrix}, \quad (2)$$

where C is the total number of channels. The study in [2] proposed that the joint instantaneous frequency $\Omega_{\mathbf{x}_+}(t)$ can be expressed as

$$\Omega_{\mathbf{x}_+}(t) = \frac{\sum_{c=1}^C \zeta_c^2(t)\omega_c(t)}{\sum_{c=1}^C \zeta_c^2(t)}, \quad (3)$$

where $\omega_c(t)$ can be computed as the derivative of $\psi_c(t)$. According to [3], the joint analytic spectrum of $\mathbf{x}_+(t)$ is

$$S_{\mathbf{x}_+}(\omega) = \frac{1}{E_{\mathbf{x}_+}} \|\mathbf{X}_+(\omega)\|^2. \quad (4)$$

C. Adaptive Fourier Decomposition

The AFD applies a matching pursuit process to search adaptive basis $\{B_n\}_{n=1}^{\infty}$ where functions B_n is called modified

TABLE I
TABLE OF NOTATIONS

Notation	Description
$x(t), x_+(t)$	Single channel real-valued signal, corresponding single channel analytic signal
$\zeta(t), \psi(t)$	Instantaneous amplitude and phase of the single channel analytic signal
$\mathbf{x}(t), \mathbf{x}_+(t)$	Multi-channel real-valued signal, corresponding multi-channel analytic signal
$x_c(t), \hat{x}_c(t)$	Signal and corresponding reconstructed signal in the c -th channel
$\zeta_c(t), \psi_c(t), \omega_c(t)$	Instantaneous amplitude, phase and frequency in the c -th channel
$\Omega_{\mathbf{x}_+}(t), S_{\mathbf{x}_+}(\omega), \mathbf{X}_+(\omega), E_{\mathbf{x}_+}$	Joint instantaneous frequency, joint analytic spectrum, frequency representation, and total energy of \mathbf{x}_+
$B_n(e^{jt})$	Modified Blaschke product
a_n	n -th parameter of B_n
ρ_m, ϕ_l	m -th magnitude and l -th phase in the discrete searching dictionary of a_n
$e_{\{a_n\}}(e^{jt}), e_{m,l}[k]$	Evaluator at a_n and its discrete form
$E_{\{B_n\}}(a), E_{\{B_n\}}[l, m]$	Objective function of searching n -th parameters of B_n and its discrete form
$G_c(e^{jt}), G_{c,n}(e^{jt}), G_{c,n}[k]$	Initial processed analytic signal in c -th channel, processed analytic signal and its discrete form in the n -th decomposition level and the c -th channel
$A_n, A_{c,n}$	Decomposition coefficient in the n -th decomposition level for single channel signal, and decomposition coefficient in the n -th decomposition level and c -th channel of the multi-channel signal
$I_n(e^{jt}), I_n[k]$	Inner function and its discrete form in the n -th decomposition level
$r_{n,h}$	h -th parameter of I_n
α_m, γ_l	m -th amplitude and l -th phase of discrete searching dictionary of $r_{n,h}$
$E_{\{I_n,h\}}[l, m]$	Objective function of searching h -th parameter of I_n
$\mu_{n,\text{core}}(t), \omega_{n,\text{core}}(t)$	Instantaneous phase and instantaneous frequency of the basis component in the n -th decomposition level for the multi-channel core AFD
$\mu_{n,\text{unwinding}}(t), \omega_{n,\text{unwinding}}(t)$	Instantaneous phase and instantaneous frequency of the basis component in the n -th decomposition level for the multi-channel unwinding AFD
$\lambda_{c,n}(t), \lambda_c(t, \omega), \Lambda(t, \omega)$	Transient amplitude in the n -th decomposition level and c -th channel, amplitude of the time-frequency distribution in the c -th channel, and joint amplitude of the multi-channel time-frequency distribution

Blaschke products and is defined as

$$B_n(e^{jt}) = \frac{\sqrt{1 - |a_n|^2}}{1 - \bar{a}_n e^{jt}} \prod_{d=0}^{n-1} \frac{e^{jt} - a_d}{1 - \bar{a}_d e^{jt}}, \quad (5)$$

$a_n \in \mathbb{D}$ ($n = 1, 2, \dots$), $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$, and \mathbb{C} is the complex plane [27]. The system $\{B_n\}_{n=1}^{\infty}$ is orthonormal for any sequence of a_n in \mathbb{D} . Therefore, the processed signal G can

be expressed as

$$G(e^{jt}) = \sum_{n=0}^{\infty} \{A_n \cdot B_n(e^{jt})\}, \quad (6)$$

where the coefficient A_n in the n -th decomposition level can be computed by $\langle G, B_n \rangle$ [16]. By defining the reduced remainder G_n in the n -th decomposition level as

$$G_n(e^{jt}) = \left(G(e^{jt}) - \sum_{n=0}^{n-1} A_n B_n(e^{jt}) \right) \prod_{d=0}^{n-1} \frac{1 - \bar{a}_d e^{jt}}{e^{jt} - a_d}, \quad (7)$$

we have the relationship

$$A_n = \langle G, B_n \rangle = \langle G_n, \mathbf{e}_{\{a_n\}} \rangle \quad (8)$$

where $\mathbf{e}_{\{a\}}$ is called evaluator or normalized reproducing kernel at a and is defined as

$$\mathbf{e}_{\{a\}}(e^{jt}) = \frac{\sqrt{1 - |a|^2}}{1 - \bar{a}e^{jt}}. \quad (9)$$

There is a long history related to the studies on the modified Blaschke products B_n [28], [29]. It should be noticed that B_n is only determined by the sequence $\{a_d\}_{d=0}^n$. When $a_0 = a_1 = \dots = a_n = 0$, then B_n is reduced to a Fourier basis component, i.e., e^{jnt} . From this point of view, the AFD basis can be considered as the generalization of the conventional Fourier basis. In addition, let the first component a_0 be 0, decomposition components can always be MCs that only have non-negative phase derivatives [16], [27]. In this case, the decomposition components are said to have well defined instantaneous frequencies [30]. Moreover, when $a_0 = 0$, $B_0(e^{jt}) = 1$, which extracts the DC components in the processed signals. Normally, key characteristics of signals are reflected by variances of signals. Therefore, in the following contents, DC components of processed signals are removed in the pre-processing, which makes the decomposition coefficients at the decomposition level $n = 0$ be equal to 0. The related decomposition processes at $n = 0$ is considered as the pre-decomposition and will not be discussed. Then, the reduced remainder at the decomposition level $n = 1$ is $G_1(e^{jt}) = G(e^{jt}) \cdot e^{-jt}$.

Starting from the decomposition level $n = 1$, the key procedure of the AFD is to determine a suitable a_n value that can provide the fast energy convergence. The achievement of the optimality is based on the attainability of

$$a_n = \arg \max \{E_{\{B_n\}}(a) : a \in \mathbb{D}\}, \quad (10)$$

where $E_{\{B_n\}}(a)$ denotes the energy of the decomposition component $B_n(e^{jt})$ that uses a as a_n in the n -th decomposition level. The attainability is referred as maximal selection principle (MSP). In this study, two main AFD algorithms, i.e., the core AFD and the unwinding AFD, are extended to the multi-channel versions. Since the core AFD and the unwinding AFD apply different basis forms and thus different schemes to search basis components, the computations of $E_{\{B_n\}}(a)$ in them are different, which will be introduced in the detailed decomposition procedures of the multi-channel core AFD and the multi-channel unwinding AFD.

In real implementations, the optimization problem shown in (10) is solved by the exhaustive searching. First, a discrete searching dictionary containing possible values of $a \in \mathbb{D}$ is generated. Then, objective function values of a values in the

searching dictionary are evaluated. Finally, the maximum objective function value and corresponding a value can be searched. As the density of the searching dictionary increases, the accuracy of optimization results will also increase. However, the computational efficiency will be declined. In the following sections, FFT based modifications of the exhaustive searching will be discussed for the discrete multi-channel core and unwinding AFD methods, which can improve the computation efficiency of this searching process.

III. MULTI-CHANNEL ADAPTIVE FOURIER DECOMPOSITION

We attempt to propose an extended AFD to analyze multi-channel signals. To achieve the alignment of common or joint oscillations across channels, the same number of MCs with the same basis components should be applied for all channels. The core AFD is the fundamental 1D AFD method that will be first generalized to the multi-channel case in this section. The unwinding AFD applies the Nevanlinna factorization to speed up the convergence and thus can provide good performance for analyzing real-world signals [24]. As the supplement of the multi-channel core AFD, the multi-channel extension of the unwinding AFD is also derived for the multi-channel signals that have common inner functions across channels. The multi-channel extension of the cyclic AFD is not included in this study because i) the cyclic AFD requires to know the number of basis components in advance, which cannot be achieved in general for real-world applications; ii) the decomposition of the cyclic AFD is similar to that of the core AFD, which leads that the extension process of the cyclic AFD can be derived by following that of the core AFD. The final implementation of the multi-channel AFD combines the multi-channel core AFD and the multi-channel unwinding AFD. After obtaining the decomposition components of the MAFD, joint instantaneous frequencies and joint instantaneous amplitudes are also determined by computing the TTFDs across channels.

A. Multi-Channel Core Adaptive Fourier Decomposition

The core AFD is the fundamental implementation of the AFD. It uses the modified Blaschke products as basis for the decomposition [16]. Following (6), the analytic signal in the c -th channel can be expressed as

$$G_c(e^{jt}) = \sum_{n=1}^N \{A_{c,n} \cdot B_n(e^{jt})\} + R_{c,N}(e^{jt}), \quad (11)$$

where the coefficient $A_{c,n}$ of the n -th decomposition level can be computed as

$$A_{c,n} = \langle G_{c,n}(e^{jt}), \mathbf{e}_{\{a_n\}}(e^{jt}) \rangle, \quad (12)$$

where the reduced remainder $G_{c,n}(e^{jt})$ can be obtained through the recursive procedure that is

$$\begin{aligned} G_{c,n}(e^{jt}) &= R_{c,n-1}(e^{jt}) \prod_{d=1}^{n-1} \frac{1 - \bar{a}_d e^{jt}}{e^{jt} - a_d} \\ &= (G_{c,n-1}(e^{jt}) - A_{c,n-1} \cdot \mathbf{e}_{\{a_{n-1}\}}(e^{jt})) \frac{1 - \bar{a}_{n-1} e^{jt}}{e^{jt} - a_{n-1}}. \end{aligned} \quad (13)$$

In order to achieve the fast energy convergence, a_n in each decomposition level should be able to extract largest energy from the total energy of the multi-channel analytic signal $\mathbf{x}_+(t)$. In the L^2 -norm sense, the total energy at the n -th decomposition

level can be computed by

$$E_{\{B_n\}}(a_n) = \sum_{c=1}^C \left| \langle G_{c,n}(e^{jt}), \mathbf{e}_{\{a_n\}}(e^{jt}) \rangle \right|^2. \quad (14)$$

Then, by inserting $E_{\{B_n\}}(a)$ shown in (14) into (10), the optimization problem of determining a_n in the multi-channel core AFD can be obtained. Although a_n determined by (14) cannot extract the largest energy of each channel simultaneously, it can extract the largest total energy of all channels. In this study, all channels are treated equally. By considering the differences of channels, a more general form of the objective function (14) is the weighted summation of $|\langle G_{c,n}(e^{jt}), \mathbf{e}_{\{a\}}(e^{jt}) \rangle|^2$ where weights indicate importance of channels. For such general case, the existence of $\{a_n\}_{n=1}^{\infty}$ and the convergence have been proved in [31]. This study further derives the MAFD for the discrete signals

For the discrete multi-channel signal, assume that the signals in all channels are equispaced discrete signal sampled from one period of periodic continuous signals with $t \in [0, 2\pi)$, and let the sampling time $\{t_k\}_{k=1}^K$ be same for all channels where $t_k = 2\pi(k-1)/K$ and K is the total number of sampling points. The searching process of a_n requires the discrete searching dictionary generated by making amplitudes and phases of values in the searching dictionary discrete, respectively. Let the discrete searching dictionary be fixed for all decomposition levels, then the objective function (14) for the discrete signal can be represented as

$$E_{\{B_n\}}[l, m] = \sum_{c=1}^C \left| \sum_{k=1}^K \{G_{c,n}[k] \cdot \{\mathbf{e}_{m,l}[k]\}^*\} \right|^2, \quad (15)$$

where $\{\cdot\}^*$ denotes the complex conjugate, and $\mathbf{e}_{m,l}[k]$ is defined as

$$\mathbf{e}_{m,l}[k] = \mathbf{e}_{\{\rho_m e^{j\phi_l}\}}(t_k) = \frac{\sqrt{1 - \rho_m^2}}{1 - \rho_m e^{j(t_k - \phi_l)}}. \quad (16)$$

Suppose the discrete phases in the searching dictionary follow sampling time, which means $t_k = \phi_l \forall k = l$, then $E_{\{B_n\}}[l, m]$ can be represented as

$$E_{\{B_n\}}[l, m] = \sum_{c=1}^C \left| \{G_{c,n} * \mathbf{e}_{m,1}\}[l] \right|^2, \quad (17)$$

where $\{G_{c,n} * \mathbf{e}_{m,1}\}[l]$ denotes the discrete convolution of $G_{c,n}[l]$ and $\mathbf{e}_{m,1}[l]$. Note that $\phi_1 = t_1 = 0$. Applying the convolution theorem to (17), the optimization problem in (10) with the objective function (14) can be generated for discrete signals, which is

$$\begin{aligned} & \text{maximize } E_{\{B_n\}}[l, m] \\ & = \sum_{c=1}^C \left| \mathcal{F}^{-1} \{ \mathcal{F} \{ G_{c,n}[l] \} \cdot \mathcal{F} \{ \mathbf{e}_{m,1}[l] \} \} \right|^2, \quad (18) \\ & \text{subject to } 0 \leq \rho_m < 1, \end{aligned}$$

where $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ denote the discrete Fourier transform (DFT) and the inverse DFT (IDFT) respectively.

Comparing the objective function of the continuous multi-channel core AFD shown in (14) and the objective function of the discrete multi-channel core AFD shown in (18), some key points in the modification should be noticed. First, only the evaluators

$\mathbf{e}_{\{\rho_m e^{j\phi_l}\}}$ at points with $\phi_1 = 0$ are required to calculate the objective functions at all points in the searching dictionary, which can significantly reduce the time of computing evaluators. Then, the FFT can be applied to compute the DFT, which increases the efficiency of computing the objective function values. The detailed analysis of the computational complexity of the basis searching is discussed in Section V-B. The analysis shows that such modification decreases the computational complexity of the MAFD, and thus improves its practicality in real-world applications.

B. Multi-Channel Unwinding Adaptive Fourier Decomposition

In Section III-A, the multi-channel core AFD does not consider the inner function part. Suppose there are common inner functions across channels, considering these common inner functions in the basis searching can significantly increase the convergence rate. According to the Nevanlinna factorization theorem [32], an analytic signal can be expressed as the product of its inner function and its outer function parts. The outer function part does not have zeros in the unit disc. Assuming that the singular part of the inner function is trivially 1, then the inner function part is identical with the Blaschke product defined by the zeros of the objective function in \mathbb{D} [18], [24]. Assuming that the signals in all channels have common zeros, following (6) and [18], the analytic signal in c -th channel can be expressed as

$$G_c(e^{jt}) = \sum_{n=1}^N \left\{ A_{c,n} \cdot B_n(e^{jt}) \cdot \prod_{i=1}^n I_i(e^{jt}) \right\} + R_{c,N}(e^{jt}), \quad (19)$$

where $I_i(e^{jt})$ is the inner function of the reduce remainder at i -th decomposition level and is defined as

$$I_i(e^{jt}) = \prod_{h=1}^{H_i} \frac{e^{jt} - r_{i,h}}{1 - \bar{r}_{i,h} e^{jt}}, \quad (20)$$

and $A_{c,n}$ can be computed as

$$A_{c,n} = \left\langle \frac{G_{c,n}(e^{jt})}{I_n(e^{jt})}, \mathbf{e}_{\{a_n\}}(e^{jt}) \right\rangle. \quad (21)$$

The corresponding recursive procedure of computing the reduced remainder is

$$G_{c,n+1}(e^{jt}) = \left(\frac{G_{c,n}(e^{jt})}{I_n(e^{jt})} - A_{c,n} \cdot \mathbf{e}_{\{a_n\}}(e^{jt}) \right) \frac{1 - \bar{a}_n e^{jt}}{e^{jt} - a_n}. \quad (22)$$

For the common inner function I_n , the parameters $\{r_{n,h}\}_{h=1}^{H_n}$ need to satisfy

$$G_{1,n}(r) = \dots = G_{C,n}(r) = 0. \quad (23)$$

Note that it is hard to determine r in (23) because $G_{c,n}(e^{jt})$ is always unknown under normal circumstance. Hence, (23) needs to be transferred to a solvable form. According to the Cauchy formula that is

$$G(r) = \left\langle G(z), \frac{1}{1 - \bar{r}z} \right\rangle \forall G(z) \in H^2, \quad (24)$$

(23) can be represented as

$$\left\langle G_{c,n}(e^{jt}), \frac{1}{1 - \bar{r}e^{jt}} \right\rangle = 0 \forall c \in \{1, 2, \dots, C\}. \quad (25)$$

All results of (25) belong to $\{r_{n,h}\}_{h=1}^{H_n}$ and are applied to construct the inner function I_n . Then, according to (21), the total energy at the n -th decomposition level is

$$E_{\{B_n\}}(a_n) = \sum_{c=1}^C \left| \left\langle \frac{G_{c,n}(e^{jt})}{I_n(e^{jt})}, \mathbf{e}_{\{a_n\}}(e^{jt}) \right\rangle \right|^2. \quad (26)$$

Same as the multi-channel core AFD, a_n can be determined by solving (10) with the objective function (26).

Algorithm 1: Discrete MAFD.

Input:

- $\{x_c\}_{c=1}^C$: multi-channel discrete signal to be processed;
 - K : total number of sampling points;
 - $\{\rho_m\}_{m=1}^M$: discrete amplitudes of a_n in the searching dictionary;
 - $\{\alpha_m\}_{m=1}^M$: discrete amplitudes of $r_{n,h}$ in the searching dictionary.
- 1: $t_k \leftarrow \frac{2\pi(k-1)}{K}$ ($k = 1, \dots, K$);
 - 2: Generate discrete phase of a_n and $r_{n,h}$ in the searching dictionary: $\phi_l = \gamma_l = t_k$ for $k = l$;
 - 3: Convert $\{x_c\}_{c=1}^C$ to its analytic representation $\{G_c\}_{c=1}^C$ via the Hilbert transform;
 - 4: Calculate $\mathcal{F}\{\mathbf{e}_{m,1}[l]\}$ and $\mathcal{F}\{\frac{1}{1-\alpha_m e^{j\gamma_l}}\}$ ($m = 1, \dots, M$);
 - 5: Initialization of determining a_n : $n \leftarrow 1$ and $G_{c,1}[k] \leftarrow G_c[k] \cdot e^{-jt_k}$ ($c = 1, 2, \dots, C$);
 - 6: **loop**
 - 7: Initialization of determining $r_{n,h}$:
 $Y_{c,n,1}[l] \leftarrow G_{c,n}[l]$, and $h \leftarrow 1$;
 - 8: **repeat**
 - 9: Search the suitable $(l_{n,h}, m_{n,h})$ by solving (28);
 - 10: **if** $\exists(l_{n,h}, m_{n,h})$ **then**
 - 11: $r_{n,h} \leftarrow \alpha_{m_{n,h}} e^{j\gamma_{l_{n,h}}}$;
 - 12: $h \leftarrow h + 1$;
 - 13: $Y_{c,n,h}[l] \leftarrow G_{c,n}[l] \cdot \prod_{i=1}^{h-1} \frac{1-\bar{r}_{n,i} e^{j\gamma_l}}{e^{j\gamma_l} - r_{n,i}}$;
 - 14: **end if**
 - 15: **until** $\nexists(l_{n,h}, m_{n,h})$
 - 16: $H_n \leftarrow$ total number of elements in $\{r_{n,h}\}$;
 - 17: **if** $H_n = 0$ **then**
 - 18: $I_n[k] \leftarrow 1$;
 - 19: **else**
 - 20: $I_n[k] \leftarrow \prod_{h=1}^{H_n} \frac{e^{jt_k} - r_{n,h}}{1 - \bar{r}_{n,h} e^{jt_k}}$;
 - 21: **end if**
 - 22: Search the suitable (l_n, m_n) by solving (32);
 - 23: $a_n \leftarrow \rho_{m_n} e^{j\phi_{l_n}}$;
 - 24: $\mathbf{e}_{\{a_n\}}[l] \leftarrow \frac{\sqrt{1-|a_n|^2}}{1-\bar{a}_n e^{j\phi_l}}$;
 - 25: $A_{c,n} \leftarrow \sum_{k=1}^K \left\{ \frac{G_{c,n}[k]}{I_n[k]} \cdot \{\mathbf{e}_{\{a_n\}}[k]\}^* \right\}$;
 - 26: $n \leftarrow n + 1$;
 - 27: $G_{c,n}[l] \leftarrow \left\{ \frac{G_{c,n-1}[l]}{I_{n-1}[l]} - A_{c,n-1} \cdot \mathbf{e}_{\{a_{n-1}\}}[l] \right\} \frac{1-\bar{a}_{n-1} e^{j\phi_l}}{e^{j\phi_l} - a_{n-1}}$;
 - 28: **end loop**
- Output:**
- $\{r_{n,h}\}$: Parameters of inner functions;
 - $\{a_n\}$: Parameters of outer functions;
 - $\{A_{c,n}\}$: Coefficients of MCs.
-

Similarly to the multi-channel core AFD, the computational efficiency of the multi-channel unwinding AFD also can be improved by modifying the objective functions of searching basis parameters and using the FFT. First, parameters $\{r_{n,h}\}_{h=1}^{H_n}$ of the inner functions need to be found by solving (25). Same as the searching process of a_n , $\{r_{n,h}\}_{h=1}^{H_n}$ can be searched from a discrete searching dictionary. In addition, the discrete searching dictionary can be constructed by the discrete amplitudes and the discrete phases of r . In the real computations, due to the computational error and the uncompleted discrete searching dictionary that cannot contain all points of $r \in \mathbb{D}$, it is hard to find a value of r that exactly makes the reduced remainders equal to zero in all channels simultaneously. However, from the searching dictionary, we can find points that are very close to the real solutions of (25) as $\{r_{n,h}\}_{h=1}^{H_n}$. These estimated points $\{r_{n,h}\}_{h=1}^{H_n}$ should satisfy

$$\sum_{c=1}^C \left| \left\langle G_{c,n}(e^{jt}), \frac{1}{1-\bar{r}_{n,h} e^{jt}} \right\rangle \right|^2 < \epsilon \quad \forall h \in \{1, 2, \dots, H_n\}, \quad (27)$$

where ϵ is a small positive number. Since the total number H_n is uncertainty, all solutions can be searched iteratively. In each iteration step, one value in $\{r_{n,h}\}_{h=1}^{H_n}$ can be determined by solving

$$\begin{aligned} & \text{minimize } E_{\{I_{n,h}\}}[l, m], \\ & \text{subject to } 0 \leq \alpha_m < 1 \text{ and } E_{\{I_{n,h}\}}[l, m] < \epsilon, \end{aligned} \quad (28)$$

where the objective function is

$$E_{\{I_{n,h}\}}[l, m] = \sum_{c=1}^C \left| \sum_{k=1}^K \left\{ Y_{c,n,h}[k] \cdot \left\{ \frac{1}{1-\alpha_m e^{j(t_k-\gamma_l)}} \right\}^* \right\} \right|^2 \quad (29)$$

and

$$Y_{c,n,h}[k] = G_{c,n}[k] \cdot \prod_{i=1}^{h-1} \frac{1-\bar{r}_{n,i} e^{jt_k}}{e^{jt_k} - r_{n,i}}. \quad (30)$$

Suppose $t_k = \gamma_l \quad \forall k = l$, then, by applying the convolution theorem, the objective function in (28) for discrete data can be represented as

$$E_{\{I_{n,h}\}}[l, m] = \sum_{c=1}^C \left| \mathcal{F}^{-1} \left\{ \mathcal{F} \{Y_{c,n,h}[l]\} \cdot \mathcal{F} \left\{ \frac{1}{1-\alpha_m e^{j\gamma_l}} \right\} \right\} \right|^2. \quad (31)$$

After obtaining all possible values of $\{r_{n,h}\}_{h=1}^{H_n}$, the inner function I_n can be constructed. The searching process of a_n in the discrete multi-channel unwinding AFD is similar to that in the discrete multi-channel core AFD, which can be expressed as

$$\begin{aligned} & \text{maximize } E_{\{B_n\}}[l, m] \\ & = \sum_{c=1}^C \left| \mathcal{F}^{-1} \left\{ \mathcal{F} \left\{ \frac{G_{c,n}[l]}{I_n[l]} \right\} \cdot \mathcal{F} \{ \mathbf{e}_{m,1}[l] \} \right\} \right|^2, \end{aligned} \quad (32)$$

subject to $0 \leq \rho_m < 1$.

It should be noticed that, for an arbitrary set of time series, the common inner functions for all channels rarely exist. However, once the common zeros exist, considering the inner function part can exactly increase the convergence rate. The

multi-channel unwinding AFD works as a supplement of the multi-channel core AFD. In one decomposition level, we first try to search parameters of the inner function by solving (28). Suppose that the common inner function cannot be found, the basis component of the multi-channel core AFD will be applied by setting the inner function part as 1. Otherwise, the common inner function will be included in the basis component. Combining the multi-channel core AFD and the multi-channel unwinding AFD, the pseudocode is shown in Algorithm 1. In this implementation, M is the number of amplitudes in the discrete searching dictionary, which controls the trade-off between the computational accuracy and efficiency. Specifically, the larger M leads to the more accurate optimization results of a_n and $r_{n,h}$ but longer computational time and larger computational resources. Therefore, M needs to be carefully pre-defined by users. The MAFD code is available at <https://toolbox-for-adaptive-fourier-decomposition.readthedocs.io>.

C. Time-Frequency Analysis Based on Multi-Channel Adaptive Fourier Decomposition

After the decomposition of the MAFD, the processed multi-channel analytic signal can be expressed as

$$\begin{aligned} \mathbf{x}_+(t) &= \begin{bmatrix} A_{1,1}e^{j\mu_1(t)} + \dots + A_{1,N}e^{j\mu_N(t)} \\ A_{2,1}e^{j\mu_1(t)} + \dots + A_{2,N}e^{j\mu_N(t)} \\ \vdots \\ A_{C,1}e^{j\mu_1(t)} + \dots + A_{C,N}e^{j\mu_N(t)} \end{bmatrix} \\ &= \begin{bmatrix} A_{1,1}, \dots, A_{1,N} \\ A_{2,1}, \dots, A_{2,N} \\ \vdots \\ A_{C,1}, \dots, A_{C,N} \end{bmatrix} \times \begin{bmatrix} e^{j\mu_1(t)} \\ e^{j\mu_2(t)} \\ \vdots \\ e^{j\mu_N(t)} \end{bmatrix}, \quad (33) \end{aligned}$$

where $e^{j\mu_n(t)}$ presents adaptive basis components searched from the MAFD. These basis components have different forms for the multi-channel core AFD and the multi-channel unwinding AFD, which leads different transient frequencies.

For the multi-channel core AFD, the n -th basis component is $e^{j\mu_{n,\text{core}}(t)} = B_n(e^{jt})$. Assuming that the processed signal is under one period $t \in [0, 2\pi)$, the transient frequency of $B_n(e^{jt})$ has been discussed in [15], which is

$$\begin{aligned} \omega_{n,\text{core}}(t) &= \frac{d\mu_{n,\text{core}}(t)}{dt} \\ &= \frac{|a_n| \cos(t - \angle a_n) - |a_n|^2}{1 - 2|a_n| \cos(t - \angle a_n) + |a_n|^2} \\ &\quad + \sum_{d=0}^{n-1} \frac{1 - |a_d|^2}{1 - 2|a_d| \cos(t - \angle a_d) + |a_d|^2}. \quad (34) \end{aligned}$$

For the multi-channel unwinding AFD, the n -th basis component is $e^{j\mu_{n,\text{unwinding}}(t)} = B_n(e^{jt}) \prod_{i=1}^n I_i(e^{jt})$. The transient frequency of the n -th decomposition component obtained from the multi-channel unwinding AFD is

$$\begin{aligned} \omega_{n,\text{unwinding}}(t) &= \frac{d\mu_{n,\text{unwinding}}(t)}{dt} \\ &= \frac{|a_n| \cos(t - \angle a_n) - |a_n|^2}{1 - 2|a_n| \cos(t - \angle a_n) + |a_n|^2} \end{aligned}$$

$$\begin{aligned} &+ \sum_{d=0}^{n-1} \frac{1 - |a_d|^2}{1 - 2|a_d| \cos(t - \angle a_d) + |a_d|^2} \\ &+ \sum_{i=1}^n \sum_{h=1}^{H_i} \frac{1 - |r_{i,h}|^2}{1 - 2|r_{i,h}| \cos(t - \angle r_{i,h}) + |r_{i,h}|^2}. \quad (35) \end{aligned}$$

The transient frequency has three key characteristics. Firstly, the transient frequencies in one decomposition level are same for all channels because the MAFD applies the same basis components for all channels. Secondly, the transient frequencies shown in (34) and (35) are always positive, which is inherited from the characteristics of the AFD and has been discussed in [27], [29], [33], [34]. The proofs for the proposed MAFD are provided in Section S.I of the supplementary material. Thirdly, owing to the characteristics of the TM system that is

$$\omega_n(t) < \omega_{n+1}(t), \quad (36)$$

there are not overlapping frequency components between different decomposition levels in each time point, which has been discussed in [15], [16] and is proofed in Section S.II of the supplementary material.

The transient amplitudes of the multi-channel core AFD and the multi-channel unwinding AFD are same because the magnitudes of the inner functions in the multi-channel unwinding AFD are 1. As reported in [15], the transient amplitude of the time-frequency distrution at the c -th channel is

$$\lambda_c(t, \omega) = \sum_{n=1}^N \lambda_{c,n}(t) \cdot \delta(\omega - \omega_n(t)), \quad (37)$$

where $\delta(\omega)$ is the unit impulse function, and

$$\lambda_{c,n}(t) = |A_{c,n}| \frac{\sqrt{1 - |a_n|^2}}{1 - |a_n| \cos(t - \angle a_n) + |a_n|^2}. \quad (38)$$

Then, according to (4), the joint transient amplitude of the multi-channel TTFD can be estimated as

$$\lambda(t, \omega) = \sqrt{\sum_{c=1}^C \lambda_c^2(t, \omega)}. \quad (39)$$

IV. SIMULATIONS AND EXAMPLES OF REAL-WORLD APPLICATIONS

The discrete MAFD shown in Algorithm 1 uses the multi-channel core AFD as the foundation and the multi-channel unwinding AFD as the supplement, which considers multi-channel signals with and without common inner functions. Therefore, this implementation is applied for the following evaluations with both synthetic and real-world signals. First, the key characteristics of the MAFD are illustrated through simulations. Then, the potential performance of the MAFD is verified by two examples of real-world applications.

A. Mode-Alignment Property

The alignment of common or joint oscillations across channels in a single decomposition level is a fundamental and important property, termed as mode-alignment, for ensuring coherent multivariate signal analysis in many practical applications, such as fusion, denoising, and classification [12], [35]. To ensure that

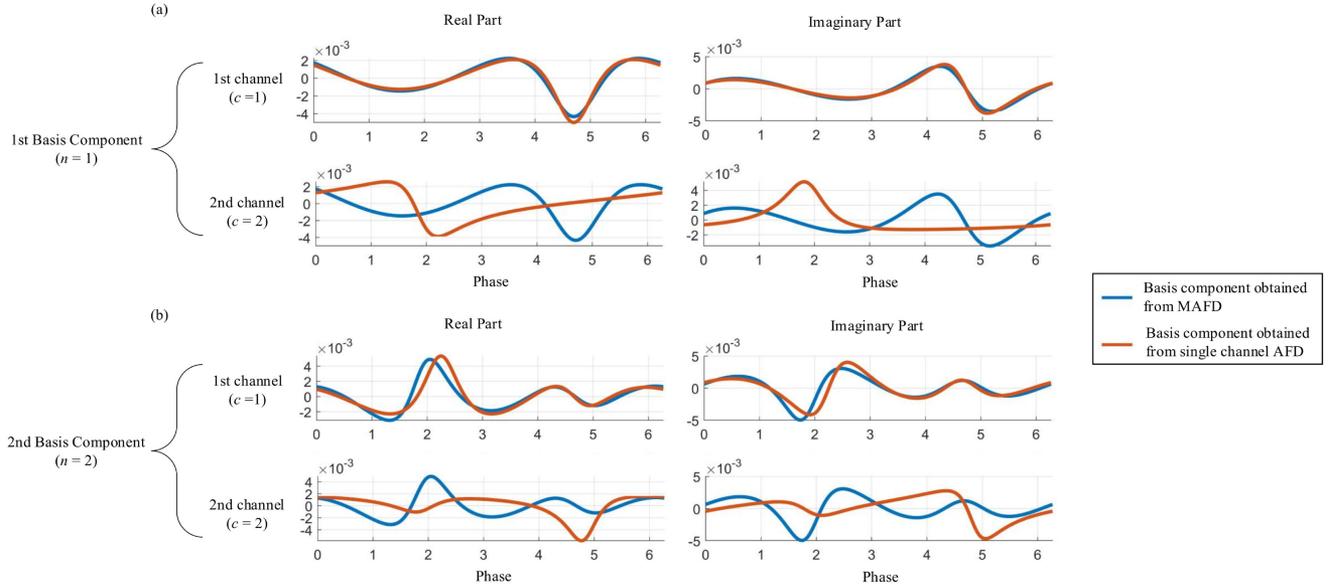


Fig. 1. Comparisons of basis components obtained from the MAFD and the single channel AFD. (a) Basis components obtained in the first decomposition level. (b) Basis components obtained in the second decomposition level.

the decomposition components can achieve the alignment of common or joint oscillations across different channels in one decomposition level, the proposed MAFD searches the adaptive common basis components for all channels. To illustrate this ability of the MAFD, we constructed a synthetic multi-channel signal that is

$$\mathbf{x}_+(t) = \begin{bmatrix} \eta_{1,1} \frac{\kappa_{1,1} e^{4j(t-\Phi_1)} + \kappa_{1,2} e^{3j(t-\Phi_1)}}{(1-\kappa_{1,3} e^{j(t-\Phi_1)})(1-\kappa_{1,4} e^{j(t-\Phi_1)})} \\ + \eta_{1,2} \frac{\kappa_{2,1} e^{4j(t-\Phi_2)} + \kappa_{2,2} e^{3j(t-\Phi_2)}}{(1-\kappa_{2,3} e^{j(t-\Phi_2)})(1-\kappa_{2,4} e^{j(t-\Phi_2)})} \\ \eta_{2,1} \frac{\kappa_{1,1} e^{4j(t-\Phi_1)} + \kappa_{1,2} e^{3j(t-\Phi_1)}}{(1-\kappa_{1,3} e^{j(t-\Phi_1)})(1-\kappa_{1,4} e^{j(t-\Phi_1)})} \\ + \eta_{2,2} \frac{\kappa_{2,1} e^{4j(t-\Phi_2)} + \kappa_{2,2} e^{3j(t-\Phi_2)}}{(1-\kappa_{2,3} e^{j(t-\Phi_2)})(1-\kappa_{2,4} e^{j(t-\Phi_2)})} \end{bmatrix}. \quad (40)$$

The similar analytic signal is also used in [16]. The basic components are common to both two channels with different summation weights $\eta_{c,n}$. Parameters of (40) are randomly generated, which are

$$\{\eta_{c,n}\} = \begin{bmatrix} 0.134 & 0.621 \\ 0.529 & 0.513 \end{bmatrix}, \quad \{\Phi_n\} = [1.882 \ 4.818],$$

and

$$\{\kappa_{n,d}\} = \begin{bmatrix} 0.435 & 0.549 & 0.420 & 0.204 \\ 0.025 & 0.435 & 0.330 & 0.619 \end{bmatrix}.$$

The single channel unwinding AFD and the proposed MAFD are applied to decompose the constructed synthetic multi-channel signal. The single channel unwinding AFD decomposes signals channel by channel while the MAFD decomposes them simultaneously. Fig. 1 shows the comparisons of basis components obtained from the single channel unwinding AFD and the MAFD. The basis components obtained from the MAFD in each decomposition level are aligned for all channels. On the contrary, the single channel unwinding AFD cannot provide the aligned basis components for all channels because the energy distributions are different in different channels.

TABLE II
RECONSTRUCTION ERROR OF THE PROPOSED MAFD, THE MNCMD, THE MEMD AND THE MWT FOR THE SYNTHETIC MULTI-CHANNEL SIGNAL

Decomposition level	1	2	3	4
MNCMD	207.999	179.893	140.665	157.713
MVMD	239.667	79.645	24.079	0.704
MEMD	180.851	160.002	0	0
MWT	398.048	383.896	372.570	363.769
MAFD	122.842	59.142	26.612	10.880

B. Fast Energy Convergence

Compared to other non-empirical decomposition methods, the fast energy convergence is one of key characteristics of the AFD owing to its adaptive basis, which is inherited by the proposed MAFD. This property of the MAFD is evaluated by the synthetic multi-channel signal applied in Section IV-A. Fig. 2 shows the comparisons of the original signals and the reconstructed signals obtained from the MAFD and the MWT at the first 10 decomposition levels. Table II shows the comparisons of the reconstruction error between empirical decomposition methods including the MNCMD, the MVMD, the MEMD, the MWT and the proposed MAFD in the first 4 decomposition levels.

To quantitatively compare the energy convergences, the relative energy error (REE) is applied, which is defined as

$$\text{REE} = \frac{\sum_{c=1}^C \left\{ \int_{t \in \mathbb{W}} |\hat{x}_c(t) - x_c(t)| \right\}}{\sum_{c=1}^C \left\{ \int_{t \in \mathbb{W}} |x_c(t)| \right\}}, \quad (41)$$

where C is the total channel number, and \mathbb{W} is the focused time window. The energy convergence rate of the MAFD is compared with those of the MWT, the MNCMD, the MVMD, and the MEMD, which is illustrated in Fig. 3.

The proposed MAFD and the MWT are both non-empirical decomposition methods. Observe that, compared with the MWT, the MAFD uses less decomposition components to reconstruct

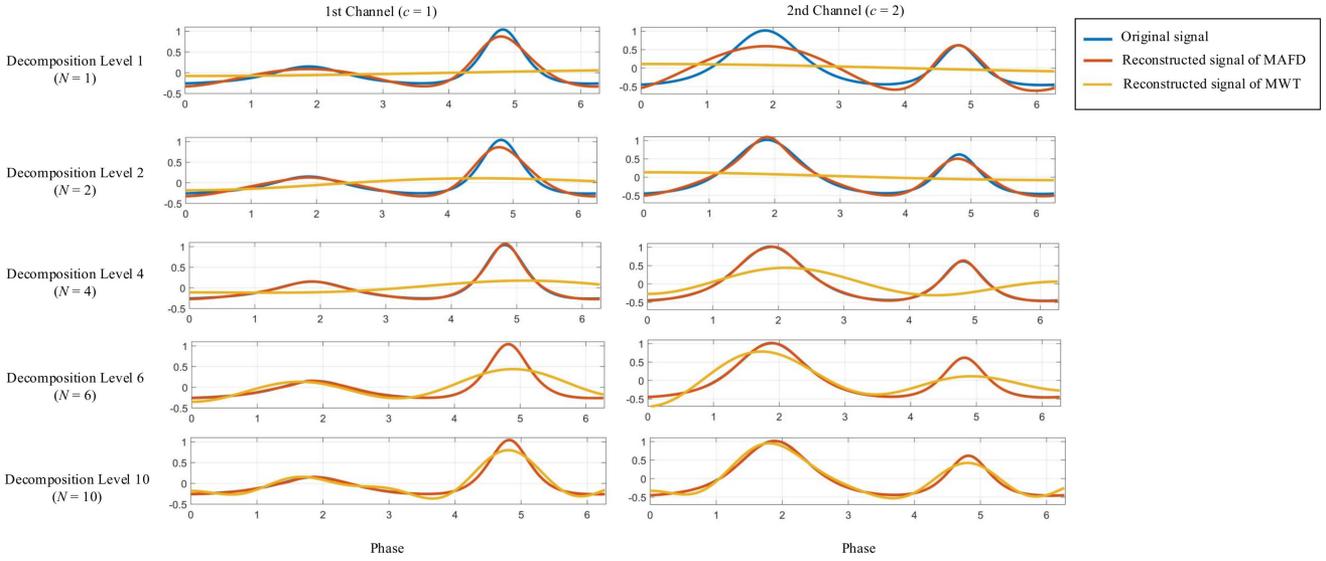


Fig. 2. Comparisons of the real parts of the reconstructed signals obtained from the MAFD and the MWT in different decomposition levels for the synthetic multi-channel signal.

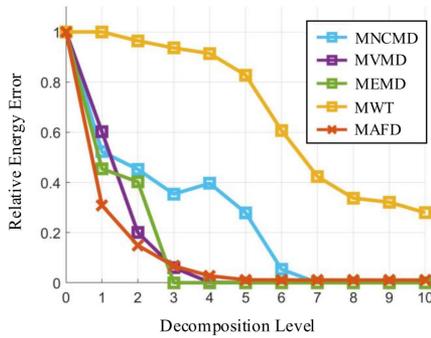


Fig. 3. Comparison of relative energy error of the proposed MAFD, the MNCMD, the MVMD, the MEMD and the MWT for the synthetic multi-channel signal.

the original signals. The MNCMD, the MVMD, and the MEMD are all empirical decomposition methods and thus normally have faster convergence than the non-empirical decomposition methods. However, the MAFD can provide the similar or even better performance compared with empirical methods. In the 3rd decomposition level, the proposed MAFD and the empirical decomposition methods already recover most energy of the original signals.

The real-world signals are normally more complex than the synthetic multi-channel signal shown in (40). To verify the energy convergence of the proposed MAFD for real-world signals, the real electrocardiogram (ECG) signals are applied. These ECG signals are from the MIT-BIH Arrhythmia Database [36], [37]. According to labels in the database, four types of ECG beats are extracted. They are the supraventricular ectopic beats, ventricular ectopic beats, fusing normal and ventricular ectopic beats, and normal beats. For each type, 100 beats are collected and combined together as a multi-channel signal to verify the energy convergence efficiency. Fig. 4 shows that, for ECG signals, the MAFD keeps the fast energy convergence. Within the first 20 decomposition levels, the REEs converge to small

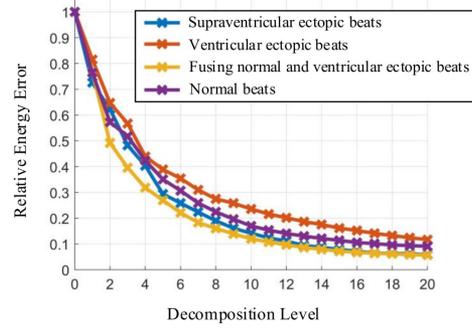


Fig. 4. Relative energy error of the MAFD for real ECG signals.

values. Fig. 5 compares the original signals and the reconstructed signals, which shows that the absolute errors between them already become very small at the 20th decomposition level.

C. Joint Time-Frequency Distribution

The joint time-frequency representation performance is evaluated by sets of sinusoidal signals, and is quantitatively measured by the localization power ratio. The localization power ratio is also used in previous studies, such as [3] and [38], and is defined as

$$B = \frac{\int \int_{(t,\omega) \in \mathbb{W}} |\text{TFD}(t, \omega)| dt d\omega}{\int \int_{(t,\omega) \notin \mathbb{W}} |\text{TFD}(t, \omega)| dt d\omega}, \quad (42)$$

where \mathbb{W} is the focused time-frequency window. The proposed MAFD is compared with the multivariate pseudo Wigner distribution (MPWD) introduced in [3], the MNCMD, the MVMD, the MEMD, the MSWT, and the MFDM. The performance is evaluated under different signal-to-noise ratios (SNRs). In each noise level, 30 independent simulation trials are carried out to avoid the influences of the stochastic process.

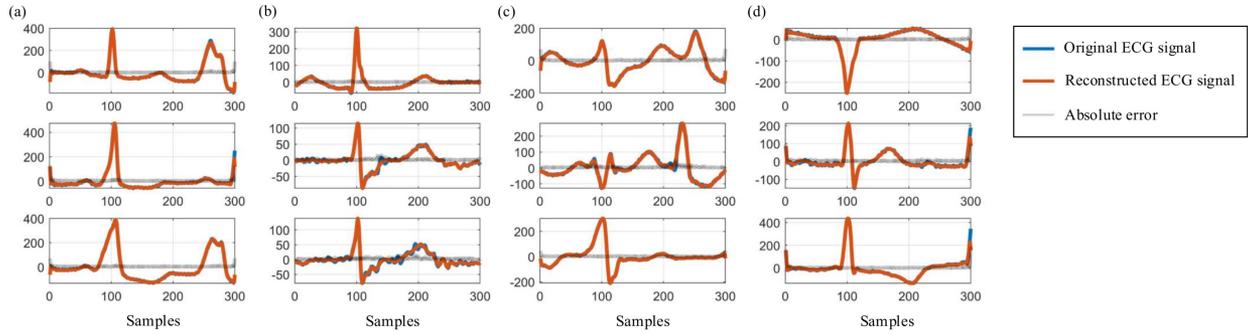


Fig. 5. Comparisons of the original ECG signals and the signals reconstructed by the MAFD at the 20th decomposition level. (a) First 3 supraventricular ectopic beats. (b) First 3 ventricular ectopic beats. (c) First 3 fusing normal and ventricular ectopic beats. (d) First 3 normal beats.

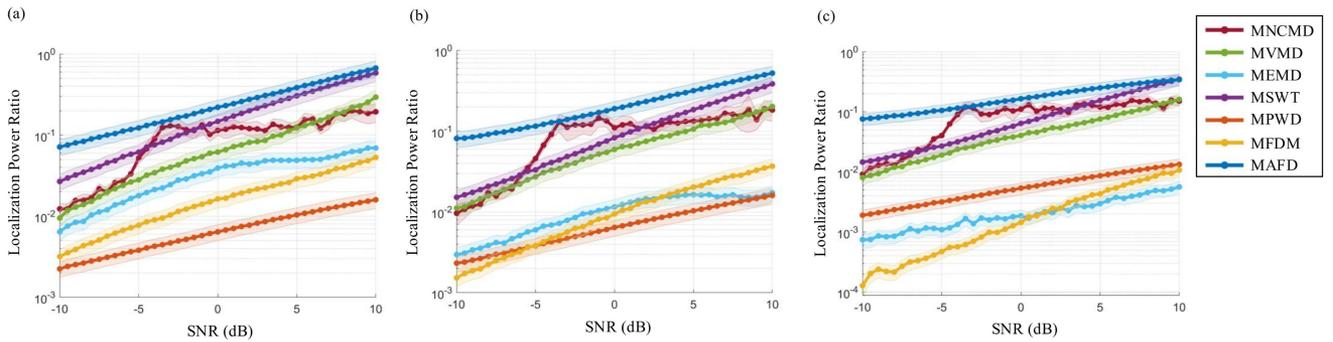


Fig. 6. Comparisons of the localization power ratio between the MAFD, the MNCMD, the MVMD, the MEMD, the MSWT, the MPWD and the MFDM for simple sinusoidal oscillations under different noise levels. (a) Frequencies of sinusoidal signals are 10 and 11 Hz. (b) Frequencies of sinusoidal signals are 40 and 41 Hz. (c) Frequencies of sinusoidal signals are 100 and 101 Hz. The dot points and the shadow areas show the averaged localization power ratios and the corresponding 95% confidence intervals.

First, a simple case of two-channel sinusoidal oscillations is considered. This two-channel signal is defined as

$$\mathbf{x}(t) = \begin{bmatrix} \sin(2\pi ft) + n_1(t) \\ \sin(2\pi(f+1)t) + n_2(t) \end{bmatrix}, \quad (43)$$

where n_c denotes the additive Gaussian white noise of the c -th channel, and f is set as 10, 40, and 100 Hz in this study. Fig. 6 shows that the proposed MAFD can provide the best time-frequency localization performance, especially for high noise levels (low SNRs). Moreover, as the SNRs increase, the localization power ratios of most methods also increase. Compared with most empirical decomposition methods, such as the MNCMD and the MEMD, the non-empirical decomposition methods, i.e., the proposed MAFD, the MSWT and the MPWD, behaves relative stable.

Then a more complex case is considered, in which the two-channel signal is constructed by the AM and the FM and corrupted with the Gaussian white noise, which is defined as

$$\mathbf{x}(t) = \begin{bmatrix} (1 + 0.5 \cos(2\pi t)) \cos(40\pi t) + n_1(t) \\ \cos(20\pi t + 3.5 \cos(2\pi t)) + n_2(t) \end{bmatrix}. \quad (44)$$

Owing to the fast energy convergence, the MAFD can provide sparse time-frequency representations, which can be observed in Fig. 7. Moreover, Fig. 8 illustrates the comparisons of the localization power ratios between the proposed MAFD, the MNCMD, the MVMD, the MEMD, the MSWT, the MPWD, and the MFDM. As the noise power decreases, the performances

of most methods increase. The proposed MAFD can provide the best performance in most cases. For low noise levels (high SNRs), since the effect of noise for the time-frequency analysis is small, the proposed MAFD, the MNCMD and the MVMD have the similar performance.

D. Examples of Real-World Applications

This section shows the potential performance of the proposed MAFD in real-world applications that are related to the electroencephalography (EEG) signals. EEG signals monitor the electrical activity of the brain and typically collected from multi-electrodes. EEG signals in multiple electrodes are normally time-varying signals and are supposed to record the same brain activity through different channels. Recently, the EEG-based brain computer interface (BCI) provides participants an effective way to communicate with the external environment without any peripheral nerves and muscles involved. Therefore, the analysis of EEG signals has gained increasing attention, especially for SSVEP signals [39], [40]. EEG signals in this study are from the SSVEP tasks in the BMI dataset [41]. This dataset contains 54 subjects' SSVEP signals. These signals are collected from 10 electrodes of the occipital region (P-7/3/z/4/8, PO-9/10, and O-1/z/2) in 200 trials of 2 sessions, i.e., the offline training phase and online test phase. Each trial data contains 4-second SSVEP signals. Signals are recorded with a sampling rate of 1000 Hz and down-sampled to 200 Hz. In the SSVEP experiments, four target SSVEP stimuli are designed to flicker at 5.45, 6.67, 8.57,

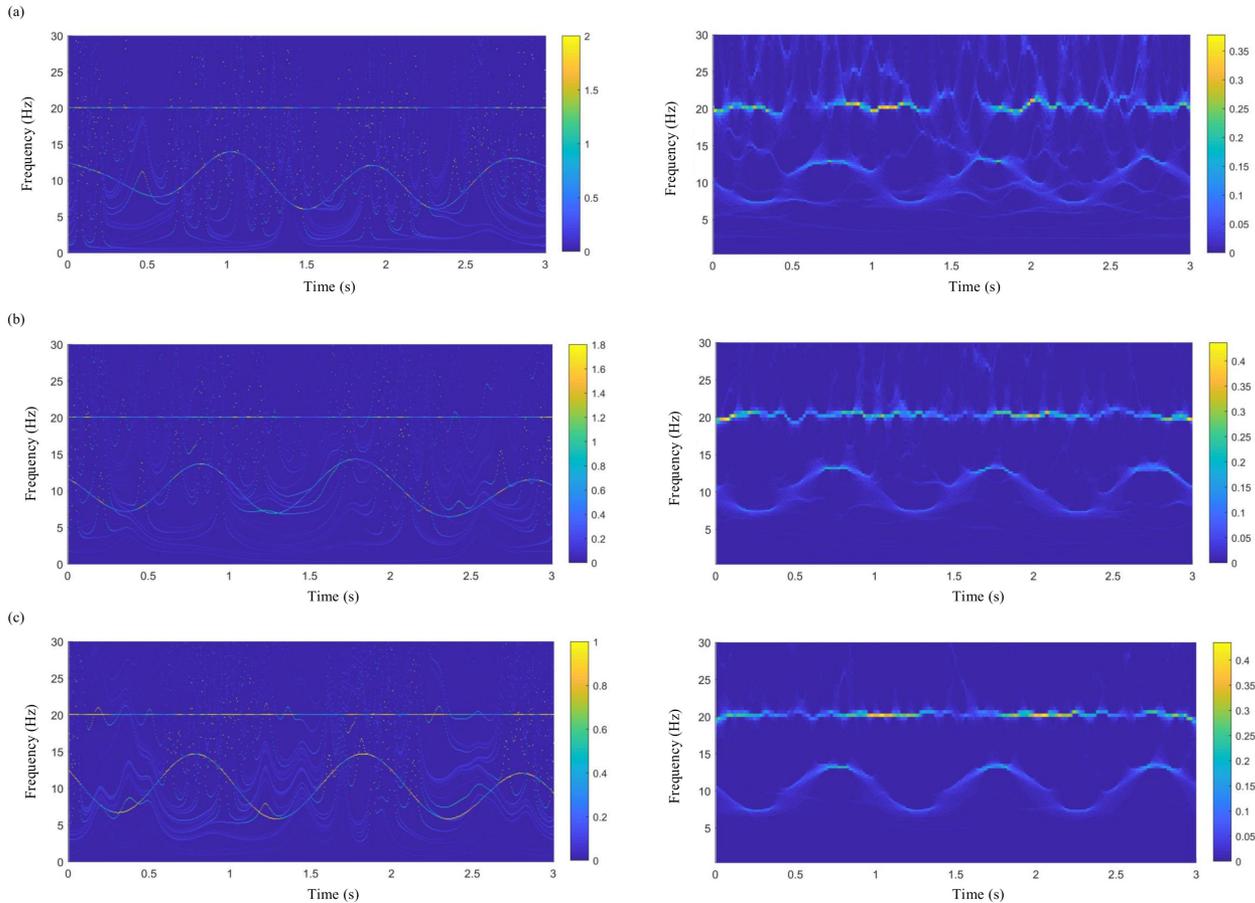


Fig. 7. Joint time-frequency representations of the proposed MAFD and the MSWT for the multi-channel signal combined with the AM and FM signals corrupted with the Gaussian white noise. Figures in the left side show the joint time-frequency representation of the MAFD. Figures in the right side show the joint time-frequency representation of the MSWT. (a) The input SNR is 0 dB. (b) The input SNR is 5 dB. (c) The input SNR is 10 dB.

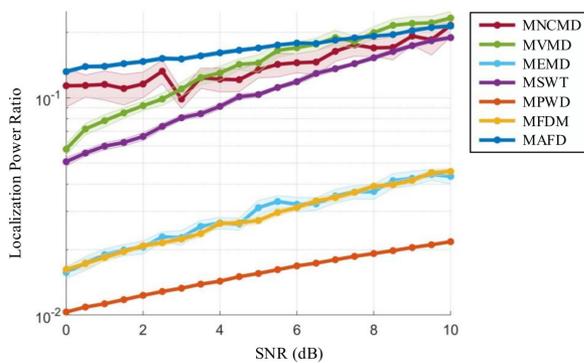


Fig. 8. Comparison of the localization power ratios between the MAFD, the MNCMD, the MVMD, the MEMD, the MSWT, the MPWD and the MFDM for the AM and FM signals under different noise levels. The dot points and the shadow areas show the averaged localization power ratios and the corresponding 95% confidence intervals.

and 12 Hz. Each stimulus is presented in 25 trials of each phases. The detailed information on these signals is available in [41]. In previous studies [21], [22], [24], [42], owing to the fast energy convergence of the AFD, the AFD performs well in the signal compression and denoising of the single-channel

signals. The following examples are to verify the signal compression and denoising performances of the proposed MAFD. Note that these examples do not focus on proposing novel compression or denoising schemes, and thus only follow the fundamental compression and denoising schemes shown in [22] and [43].

1) *EEG Signal Compression*: The proposed MAFD is compared with the MWT shown in [2]. In the comparisons, we select coefficients of first several decomposition components as compressed data. The number of decomposition components is controlled to make the REE smaller than 0.1. Under this condition, Fig. 9 as an example shows that the MAFD and the MWT both can reconstruct the original signal with very small error. The compression efficiency is measured by the compression ratio (CR), which is defined as $CR = \frac{b_{in}}{b_{out}}$ where b_{in} and b_{out} are the numbers of values in the original data and the compressed data respectively. Fig. 10 compares the CRs of the proposed MAFD and the MWT for 54 subjects individually and cumulatively. Due to individual differences, the CRs are not exactly the same for all subjects. In this comparison, the proposed method can always provide better CRs.

2) *EEG Signal Classification*: The proposed MAFD is compared with the independent component analysis (ICA). The ICA is a conventional blind source separation method and has shown good denoising performance for EEG signals [43], [44]. The

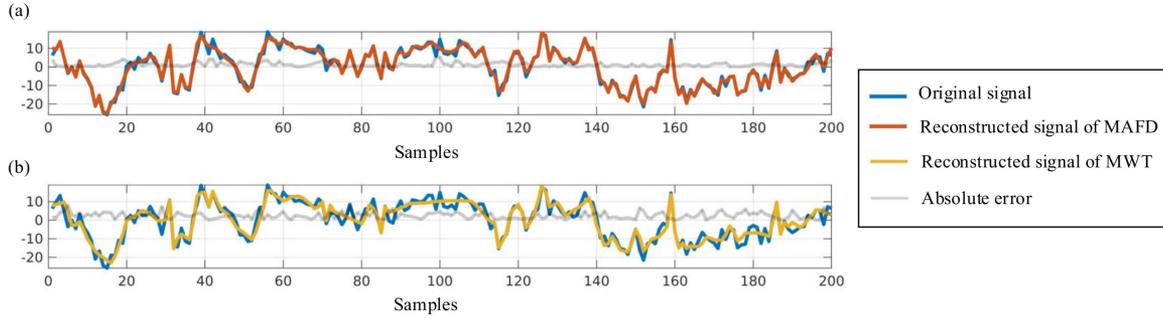


Fig. 9. Comparisons between the original EEG signal at Oz channel and the corresponding reconstructed signals. (a) Reconstructed signal of the MAFD. (b) Reconstructed signal of the MSWT.

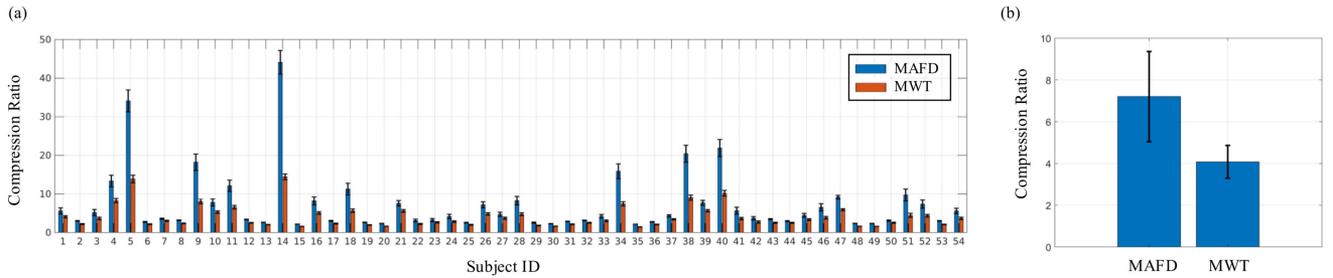


Fig. 10. Comparisons between compression ratios of the MAFD and the MWT. (a) Individual comparisons. (b) Cumulative comparisons of all subjects. The heights of bars show the averaged compression ratios. The vertical error bars denote the corresponding 95% confidence intervals.

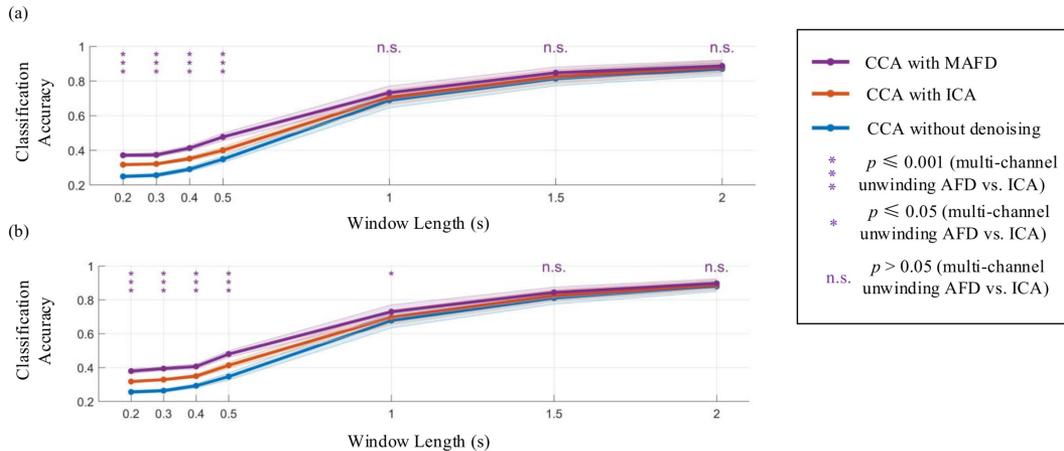


Fig. 11. Comparisons between SSVEP classification accuracies with the MAFD-based denoising, with the ICA-based denoising, and without denoising. (a) SSVEP classification accuracies of the offline training phase. (b) SSVEP classification accuracies of the online test phase. The dot points and the shadow areas show the averaged classification accuracies and the corresponding 95% confidence intervals.

denoising performance is evaluated by the classification accuracy (CA) defined as $CA = \frac{n_{\text{corr}}}{n_{\text{total}}}$ where n_{corr} and n_{total} are numbers of correct and total trials respectively. The standard canonical correlation analysis (CCA)-based SSVEP classification method is applied to categorize SSVEP signals [41], [45]. Fig. 11 shows the comparisons of the CA under different window lengths. Compared to the ICA, the proposed method can provide better CAs, especially for the small time window lengths, i.e., 0.2 s, 0.3 s, 0.4 s, and 0.5 s. In these time window lengths, Bonferroni corrected t -test results indicate that the CCA with the MAFD performs significantly better than the CCA with the ICA (all p -values are smaller than 0.001).

V. DISCUSSION

A. Comparisons of Basis Components

Unlike the single channel AFD, the MAFD generates common basis components for all channels. In each decomposition level, the single channel AFD searches the basis components by only considering the energy distribution of the single channel signal. However, the MAFD considers the energy distribution of signals in all channels. Fig. 12 shows the energy distributions of the synthetic multi-channel signal applied in Section IV-A. When the energy distributions of different channels are different as shown in Fig. 12(b) and (c), the basis components obtained

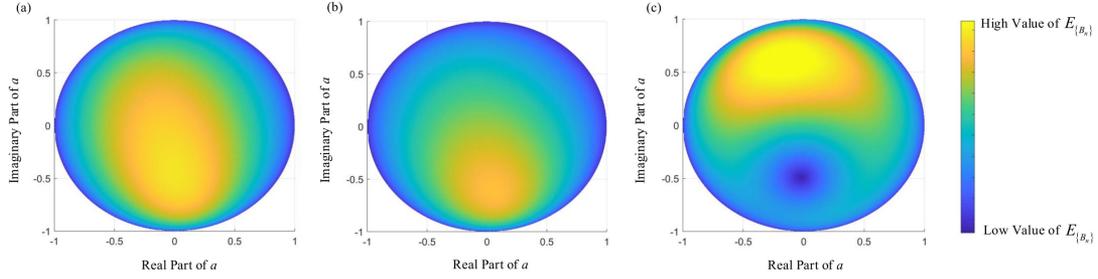


Fig. 12. $E_{\{B_n\}}(a)$ distributions of the synthetic multi-channel signal defined in Section IV-A at $n = 1$. (a) $E_{\{B_n\}}$ of the MAFD. (b) $E_{\{B_n\}}$ of the single channel AFD for the first channel. (c) $E_{\{B_n\}}$ of the single channel AFD for the second channel.

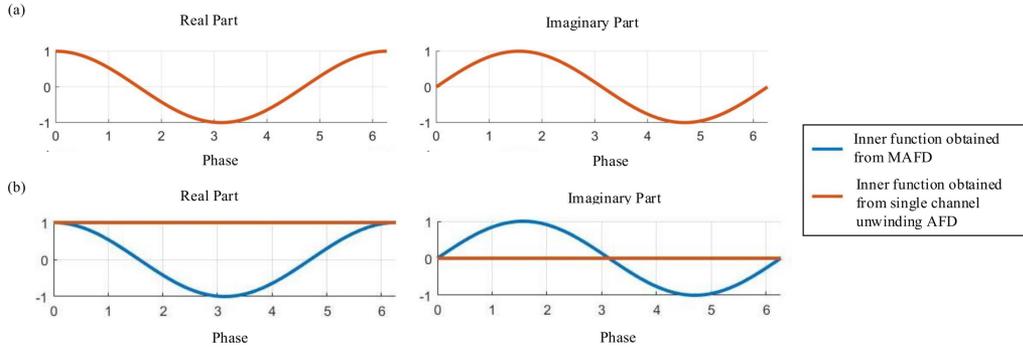


Fig. 13. Inner functions obtained from the MAFD and the single channel unwinding AFD for the synthetic multi-channel signal defined in Section IV-A at $n = 1$. (a) Inner function obtained from the first channel signal. (b) Inner function obtained from the second channel signal.

from the single channel AFD will also be different as shown in Fig. 1. In addition, since the energy distribution of the whole multi-channel signal is similar to the energy distribution of the first channel signal as shown in Fig. 12(a) and (b), the basis components obtained by the MAFD are also similar to the basis components obtained by the single channel AFD for the first channel signal as shown in Fig. 1.

The above simulations are all based on the proposed implementation of the MAFD illustrated in Algorithm 1. There are two important notices for the searching process of the inner functions. First, although we assume that signals in different channels have common inner functions to propose the multi-channel unwinding AFD, multi-channel signals that satisfy such assumption are very rare. Most multi-channel signals that do not satisfy the assumption of the multi-channel unwinding AFD should be analyzed by the multi-channel core AFD. Therefore, in the proposed implementation of the MAFD, the inner functions will not be applied to construct the basis components when the common inner function cannot be found. Second, as mentioned in Section III-B, the parameters of inner functions should satisfy (25) but they are estimated by (27) in practice. Although $r_{n,h}$ estimated by (27) can achieve $|\langle G_{c,n}, \frac{1}{1-\bar{r}_{n,h}e^{j\bar{t}}} \rangle| < \epsilon \forall c \in \{1, \dots, C\}$, there may be little differences between the estimated solutions and the real solutions of (25). Fig. 13 shows the inner functions obtained from the MAFD and the single channel unwinding AFD for the synthetic multi-channel signal applied in Section IV-A at the first decomposition level. The inner functions searched by the single channel unwinding AFD can be considered as the real inner functions of the single channel signals. It can be seen that, the second channel signal do not have the inner function but the MAFD still provide the common

inner function, which is due to the mentioned estimation error. In this situation, the decomposition based on the common basis components provided by the MAFD is not exactly factorization for the second channel signal. Although there is the estimated error, the estimated solutions still can achieve $G_{c,n}(r_{n,h}) \approx 0 \forall c \in \{1, \dots, C\}$. Therefore, such small estimated error does not affect the fast energy convergence as illustrated in Fig. 3, and thus is acceptable.

B. Computation Complexity

As mentioned in Section III, the objective functions of searching the basis parameter are reformulated and computed by the FFT for discrete signals. Such modifications can improve the computational efficiency by reducing the computation complexity of searching basis parameters, which is analyzed in this section. The computation complexity of the basis searching process is related to the total number of channels C , the total number of sampling points K , and the number of magnitude values M in the discrete searching dictionary. Suppose that the phase values in the searching dictionary follow the sampling time, then the number of phase values in the searching dictionary is also K .

First, before searching the basis parameters, all evaluators need to be generated. For the original objective function extended from the continuous MAFD and shown in (15), all evaluators $e_{m,l}[k]$ of $m \in [1, \dots, M]$, $l \in [1, \dots, K]$, and $k \in [1, \dots, K]$ need to be computed, which requires $\mathcal{O}(MK^2)$ computation steps. After the modification as shown in (18), only evaluators at one phase value are required. The computation complexity of generating evaluators $e_{m,1}[l]$ is reduced

TABLE III

COMPUTATIONAL TIMES (IN SECONDS) OF DIFFERENT NUMBER OF AMPLITUDE VALUES IN SEARCHING DICTIONARIES ($C = 4$ AND $K = 300$)

Number of amplitude values (M)	48	32	24	19
Original MAFD	3.9207	2.5491	1.8784	1.4709
FFT-MAFD	0.1046	0.0798	0.0702	0.0616
Original AFD	4.0495	2.6244	1.9751	1.5161
FFT-AFD	0.1705	0.1227	0.1025	0.0883

TABLE IV

AVERAGED COMPUTATIONAL TIMES (IN SECONDS) OF DIFFERENT CHANNEL NUMBERS ($K = 300$ AND $M = 48$)

Channel number (C)	4	6	8	16	32
Original MAFD	3.8913	5.8748	7.7320	15.4589	31.3183
FFT-MAFD	0.1051	0.1481	0.1891	0.3635	0.7269
Original AFD	4.0662	6.1946	8.2285	16.3059	32.8663
FFT-AFD	0.1700	0.2569	0.3433	0.6864	1.3842
MVMD	0.2063	0.2760	0.5081	0.9436	1.3949
MEMD	0.1222	0.1386	0.1610	0.1629	0.1897
MNCMD	1.3127	1.9094	2.5334	5.3537	10.8427
MFDM	0.0096	0.0144	0.0185	0.0383	0.0784
MWT	0.0010	0.0015	0.0019	0.0038	0.0076

to $\mathcal{O}(MK)$. Then, in each decomposition level, the objective function values need to be computed. For the original objective function extended from the continuous MAFD, the number of objective function values in each channel is same as the number of evaluators. Therefore, $\mathcal{O}(CMK^2)$ is required for computing objective function values. After the modification, the objective function values can be calculated by the FFT and the inverse FFT, which reduces the computation complexity to $\mathcal{O}(CMK(2\log_2 K + 1))$ in each decomposition level. Based on the above analysis, the computation complexity of the discrete MAFD based on the modified objective function is lower than that based on the original objective function extended from the continuous MAFD.

To quantitatively compare the computational efficiency, the computational times of different methods are measured. As mentioned in above analysis, the channel number, sampling number in each channel and the sizes of the searching dictionaries have large effects on the computational efficiency of the MAFD. For each simulation, 30 independent trials are carried out to avoid the influences of the stochastic effects. Then, computational times of 30 trials are averaged. The sizes of the searching dictionaries in the AFD and the MAFD are determined by M . The effects of M are shown in Table III where the computational times of the MAFD before the FFT based modification (original MAFD) and the MAFD after the FFT based modification (FFT-MAFD) as well as the conventional AFD before the FFT based modification (original AFD) and the conventional AFD after the FFT based modification (FFT-AFD) are compared. The effects of the channel number C and the sampling number per channel K are illustrated in Tables IV and V where the proposed MAFD is compared with the conventional AFD, the MEMD, the MVMD, the MNCMD, and the MFDM. All implementations are carried in MATLAB R2019a with an Intel Xeon Gold 6128 processor

TABLE V

COMPUTATIONAL TIMES (IN SECONDS) OF DIFFERENT SAMPLING NUMBER PER CHANNEL ($C = 4$ AND $M = 48$)

Sampling number (K)	300	600	900	1200
Original MAFD	3.9178	12.2202	25.2838	43.7299
FFT-MAFD	0.1054	0.1419	0.2460	0.3859
Original AFD	4.0513	12.2808	25.2471	41.8013
FFT-AFD	0.1702	0.1761	0.2799	0.3594
MVMD	0.2100	0.6174	0.8548	1.1069
MEMD	0.1298	0.1996	0.2687	0.3067
MNCMD	1.3395	2.5413	4.0261	5.3835
MFDM	0.0103	0.0106	0.0112	0.0116
MWT	0.0012	0.0012	0.0013	0.0013

@ 3.40 GHz. Observe that, for the MAFD and the AFD, the modifications based on the FFT can significantly reduce the computational times. As the sizes in the searching dictionaries, the channel number and the sampling number increase, the reduction effects of the FFT based modifications become more significant, which matches the above analysis of the computation complexities. When applying the AFD channel by channel, the searching dictionaries and the evaluators are required to be generated for each channel respectively, which makes the computational times of the AFD larger than the computational times of the MAFD in most cases, especially for large channel numbers. Compared with other multivariate adaptive methods including the MVMD, the MEMD, the MNCMD, and the MFDM, although the FFT based modifications can make the computational times of the MAFD become similar with or even lower than others, the MAFD still cannot provide the significant benefit on the computational efficiency. Moreover, when the numbers of points in the searching dictionaries, the channel number and the sampling number increase, there are large increments of the required computing resources and computational times for the MAFD. For example, as illustrated in Table V, when the sampling number is increased to 1200, the MAFD becomes slower than the MEMD and the MVMD. However, the MAFD has several benefits that empirical decomposition methods do not contain, which is discussed in Section V-C. Compared with the non-adaptive decomposition method, i.e., the MWT, although the MAFD is much slower, the MAFD has many advantages, such as the adaptive basis and the fast energy convergence, and thus can provide better performance on analyzing nonlinear and non-stationary signals.

C. Comparison With Empirical Decomposition Methods

In Sections IV and V-B, the proposed MAFD are compared with several novel multivariate empirical decomposition methods including the MNCMD, the MVMD, the MEMD and the MFDM. The decomposition processes of the proposed MAFD and these empirical decomposition methods are all data-driven and thus can achieve adaptive to the processed signals. However, compared with these empirical decomposition methods, the MAFD has several advantages.

Firstly, the MAFD is extended from the AFD and has the rigorous mathematical foundation. Therefore, the decomposition components provided by the MAFD have the explicit mathematical expression. This characteristic of the MAFD ensures the

stable decomposition process and can provide the mathematical support for further interpreting and analyzing decomposition components [34]. As illustrated in Fig. 6 and Fig. 8, when the noise levels decrease, the localization power ratios are expected to increase, which is also satisfied by the proposed MAFD. However, in some cases, the performances of the MNCMD and the MEMD do not follow this rule, which is hard to be explained. Secondly, the basis components of the MAFD are the modified Blaschke products that have several good characteristics, which is hard to be achieved in empirical decomposition methods [29]: i) The components in different decomposition levels are orthogonal, which avoids the mode-mixing problem in the mathematical theorem; ii) The decomposition components only contain non-negative phase derivatives, which leads the meaningful instantaneous frequency analysis. Thirdly, compared with the MVMD and the MNCMD, the number of required decomposition parameters in the MAFD is less. The key optimization problems in the MVMD and the MNCMD are based on the augmented Lagrangian functions in which the penalty parameters have large affects on the decomposition results and are normally hard to be determined. The proposed MAFD only needs the searching dictionaries of the basis parameters as shown in Algorithm 1. For general signals, how to generate such searching dictionaries have been discussed in Sections III-A and III-B.

Although there are several advantages of the MAFD, the proposed MAFD cannot totally replace the empirical decomposition methods. The proposed MAFD and the empirical methods belong two different categories. The decomposition processes of the empirical methods are directly fully derived from the processed signals and are not limited by basis. As illustrated in Fig. 3, Fig. 8 and comparisons of computational times in Section V-B, the empirical methods can provide faster computation, faster energy convergence or better time-frequency localization performance in some cases. Therefore, suppose the processed signals need to be divided into several empirical modes without further mathematical analysis, empirical methods can provide good performance. However, for some applications that requires the signal modeling and/or detailed mathematical interpretation of decomposition results, the proposed MAFD would be the good choice.

VI. CONCLUSION

We propose the MAFD, including the multi-channel core AFD and the multi-channel unwinding AFD. Compared to the conventional AFD and the conventional multi-channel signal decomposition methods, the MAFD applies the common basis components to decompose all channel signals and thus achieve the alignment of common or joint frequency components across channels in one decomposition level. Meanwhile, being a generic extension of the AFD, the proposed MAFD inherits all advantages and flaws of the original AFD.

The advantages of the AFD, i.e. the adaptive basis and the fast energy convergence as well as decomposition components that only have non-negative phase derivatives and non-overlapping frequencies across decomposition levels, are maintained in the proposed MAFD. The proposed joint TTFD also inherits good mathematical properties of the conventional TTFD. Owing to the mode-alignment property of the MAFD and advantages inherited from the AFD, simulations in this study show promising results in real world examples.

Some challenges of applying the original AFD in real applications are also observed in the proposed MAFD: i) Although the FFT based modifications can improve the computational efficiency, the computational times are still long when the numbers of total sampling points and points in the searching dictionaries are large; ii) The physical meaning of basis parameters has not been fully explored; iii) When the decomposition level is finite, there is the problem of dealing with low-energy components. Since the studies and the applications of the MAFD are still at their infancy, as more deep studies of the AFD and the MAFD, novel solutions to these shortcomings are expected. For specific applications, detailed studies focusing on combining characteristics of both processed signals and basis components to regular the selection region of the basis parameters would further improve the performance of the MAFD.

REFERENCES

- [1] D. Gabor, "Theory of communication. Part I: The analysis of information," *J. Inst. Elect. Eng.*, vol. 93, no. 26, pp. 429–441, 1946.
- [2] J. M. Lilly and S. C. Olhede, "Analysis of modulated multivariate oscillations," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 600–612, Feb. 2012.
- [3] A. Ahrabian, D. Looney, L. Stanković, and D. P. Mandic, "Synchrosqueezing-based time-frequency analysis of multivariate data," *Signal Process.*, vol. 106, pp. 331–341, 2015.
- [4] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proc. R. Soc. A. Math. Phys. Eng. Sci.*, vol. 466, no. 2117, pp. 1291–1302, 2010.
- [5] X. Lang *et al.*, "Fast multivariate empirical mode decomposition," *IEEE Access*, vol. 6, pp. 65521–65538, 2018.
- [6] N. ur Rehman and D. P. Mandic, "Filter bank property of multivariate empirical mode decomposition," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2421–2426, May 2011.
- [7] C. Park, D. Looney, N. ur Rehman, A. Ahrabian, and D. P. Mandic, "Classification of motor imagery BCI using multivariate empirical mode decomposition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 1, pp. 10–22, Jan. 2013.
- [8] H. Hao, H. L. Wang, and N. U. Rehman, "A joint framework for multivariate signal denoising using multivariate empirical mode decomposition," *Signal Process.*, vol. 135, pp. 263–273, 2017.
- [9] P. Singh, S. D. Joshi, R. K. Patney, and K. Saha, "The Fourier decomposition method for nonlinear and non-stationary time series analysis," *Proc. R. Soc. A.*, vol. 473, no. 2199, 2017, Art. no. 20160871.
- [10] P. Singh, "Novel Fourier quadrature transforms and analytic signal representations for nonlinear and non-stationary time-series analysis," *R. Soc. Open Sci.*, vol. 5, no. 11, 2018, Art. no. 181131.
- [11] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.
- [12] N. u. Rehman and H. Aftab, "Multivariate variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 67, no. 23, pp. 6039–6052, Dec. 2019.
- [13] Q. Chen, L. Xie, and H. Su, "Multivariate nonlinear chirp mode decomposition," *Signal Process.*, vol. 176, 2020, Art. no. 107667.
- [14] Q. Chen, X. Lang, L. Xie, and H. Su, "Multivariate intrinsic chirp mode decomposition," *Signal Process.*, vol. 183, 2021, Art. no. 108009.
- [15] P. Dang, T. Qian, and Y. Y. Guo, "Transient time-frequency distribution based on mono-component decompositions," *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 11, no. 3, 2013, Art. no. 1350022.
- [16] T. Qian, L. Zhang, and Z. Li, "Algorithm of adaptive Fourier decomposition," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5899–5906, Dec. 2011.
- [17] T. Qian, "Cyclic AFD algorithm for the best rational approximation," *Math. Methods Appl. Sci.*, vol. 37, no. 6, pp. 846–859, 2014.
- [18] T. Qian, "Intrinsic mono-component decomposition of functions: An advance of Fourier theory," *Math. Methods Appl. Sci.*, vol. 33, no. 7, pp. 880–891, 2010.
- [19] T. Qian, "Two-dimensional adaptive Fourier decomposition," *Math. Methods Appl. Sci.*, vol. 39, no. 10, pp. 2431–2448, 2016.
- [20] W. Mi, T. Qian, and F. Wan, "A fast adaptive model reduction method based on Takenaka-Malmquist systems," *Syst. Control Lett.*, vol. 61, no. 1, pp. 223–230, 2012.

- [21] Z. Wang, F. Wan, C. M. Wong, and L. Zhang, "Adaptive Fourier decomposition based ECG denoising," *Comput. Biol. Med.*, vol. 77, pp. 195–205, 2016.
- [22] J. Ma, T. Zhang, and M. Dong, "A novel ECG data compression method using adaptive Fourier decomposition with security guarantee in e-health applications," *IEEE J. Biomed. Health. Informat.*, vol. 19, no. 3, pp. 986–994, May 2015.
- [23] Q. Chen, T. Qian, Y. Li, W. Mai, and X. Zhang, "Adaptive Fourier tester for statistical estimation," *Math. Methods Appl. Sci.*, vol. 39, no. 12, pp. 3478–3495, 2016.
- [24] C. Tan, L. Zhang, and H.-t. Wu, "A novel Blaschke unwinding adaptive Fourier decomposition-based signal compression algorithm with application on ECG signals," *IEEE J. Biomed. Health. Informat.*, vol. 23, no. 2, pp. 672–682, Mar. 2019.
- [25] Y. Li, L. Zhang, and T. Qian, "2D partial unwinding—a novel non-linear phase decomposition of images," *IEEE Trans. Image Process.*, vol. 28, no. 10, pp. 4762–4773, Oct. 2019.
- [26] Y. Gao, T. Qian, V. Temlyakov, and L.-f. Cao, "Aspects of 2D-adaptive Fourier decompositions," 2017. [Online]. Available: [arXiv:1710.09277](https://arxiv.org/abs/1710.09277).
- [27] T. Qian and Y.-B. Wang, "Adaptive Fourier series—A variation of greedy algorithm," *Adv. Comput. Math.*, vol. 34, no. 3, pp. 279–293, 2011.
- [28] P. L. Butzer and R. J. Nessel, *Fourier Analysis and Approximation*. Basel, Switzerland: Birkhauser, 1971.
- [29] T. Qian, "Mono-components for decomposition of signals," *Math. Meth. Appl. Sci.*, vol. 29, no. 10, pp. 1187–1198, 2006.
- [30] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ, USA: Prentice Hall PTR, 1995.
- [31] T. Qian, "Sparse representations of random signals," *Math. Meth. Appl. Sci.*, to be published, doi: [10.1002/mma.8033](https://doi.org/10.1002/mma.8033).
- [32] J. Garnett, *Bounded Analytic Functions*. New York, NY, USA: Springer, 2007.
- [33] T. Qian, Q. Chen, and L. Li, "Analytic unit quadrature signals with nonlinear phase," *Physica D: Nonlinear Phenomena*, vol. 203, no. 1–2, pp. 80–87, 2005.
- [34] T. Qian, L. Zhang, and H. Li, "Mono-components vs IMFs in signal decomposition," *Int. J. Wavelets Multiresolution Inf. Process.*, vol. 6, no. 3, pp. 353–374, 2008.
- [35] D. Looney and D. Mandic, "Multiscale image fusion using complex extensions of EMD," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1626–1630, Apr. 2009.
- [36] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e 215–e220, 2000.
- [37] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, May/June 2001.
- [38] I. Djurović and L. Stanković, "Time-frequency representation based on the reassigned S-method," *Signal Process.*, vol. 77, no. 1, pp. 115–120, 1999.
- [39] S. Gao, Y. Wang, X. Gao, and B. Hong, "Visual and auditory brain computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1436–1447, May 2014.
- [40] C. M. Wong, B. Wang, Z. Wang, K. F. Lao, A. Rosa, and F. Wan, "Spatial filtering in SSVEP-based BCIs: Unified framework and new improvements," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 11, pp. 3057–3072, Nov. 2020.
- [41] M.-H. Lee, O.-y. Kwon, Y.-J. Kim, H.-k. Kim, Y.-E. Lee, J. Williamson, S. Fazli, and S.-w. Lee, "EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy," *Gigascience*, vol. 8, no. 5, pp. 1–13, 2019.
- [42] Z. Wang, C. M. Wong, and F. Wan, "Adaptive Fourier decomposition based R-peak detection for noisy ECG signals," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2017, pp. 3501–3504.
- [43] S. P. Fitzgibbon, D. M. W. Powers, K. J. Pope, and C. R. Clark, "Removal of EEG noise and artifact using blind source separation," *J. Clin. Neurophysiol.*, vol. 24, no. 3, pp. 232–243, 2007.
- [44] B. W. McMenamin *et al.*, "Validation of ICA-based myogenic artifact correction for scalp and source-localized EEG," *Neuroimage*, vol. 49, no. 3, pp. 2416–2432, 2010.
- [45] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2610–2614, Dec. 2006.



Ze Wang (Student Member, IEEE) received the B.S. and M.Sc. degrees in electrical and computer engineering, in 2014 and 2017, respectively, from the University of Macau, Macau, China, where he is currently working toward the Ph.D. degree. His research interests include biomedical signal processing, brain-computer interfaces, and machine learning.



Chi Man Wong received the B.S. degree from Jinan University, Guangzhou, China, in 2006, and the M.Sc. and Ph.D. degrees from the University of Macau, Macau, China, in 2011 and 2021, respectively. He is currently a Postdoctoral Fellow with the University of Macau. His research interests include brain-computer interfaces, biomedical signal processing, and machine learning.



Agostinho Rosa received the M.Sc. and the Ph.D. degrees in electrotechnical engineering from Instituto Superior Tecnico, the University of Lisbon, Lisbon, Portugal, in 1984 and 1991, respectively. He is currently an Associate Professor with the Bioengineering Department, University of Lisbon. His research interests include signal processing, sleep research, neuromodulation, and bio-inspired optimization.



Tao Qian received the Ph.D. degree in harmonic analysis from Peking University, Beijing, China, in 1984. From 1984 to 1986, he was with the Institute of Systems Science, Chinese Academy of Sciences, Beijing, China. He was also a Research Fellow with Macquarie University, Sydney, NSW, Australia, and held a Postdoctoral position with the Flinders University of South Australia, Adelaide, SA, Australia, until 1992, under the supervision of A. McIntosh and G. Gaudry. From 1992 to 2000, he was a Lecturer with New England University, Armidale NSW, Australia. Since 2000, he has been an Associate Professor with the University of Macau, Macau, China. He received the Full Professorship in 2003. From 2005 to 2011, he was the Head of the Department of Mathematics. From 2013 to 2018, he was appointed as Distinguished Professor of the University of Macau. Since 2019, he has been a Full Professor with the Macau University of Science and Technology, Macau, China. His research interests include pure analysis, applied and computational harmonic analysis, complex and Clifford analysis, signal analysis, and related areas.



Feng Wan (Senior Member, IEEE) received the Ph.D. degree in electrical and electronic engineering from The Hong Kong University of Science and Technology, Hong Kong, in 2002. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, Faculty of Science and Technology, and a Primary Faculty Member of the Centre for Cognitive and Brain Sciences and the Centre of Artificial Intelligence and Robotics, Institute of Collaborative Innovation, University of Macau, Macau, China. His research interests include biomedical signal processing, brain-computer interfaces, neurofeedback training, computational intelligence, and intelligent control.