

Quaternion Singular Spectrum Analysis of Electroencephalogram with Application in Sleep Analysis

Shirin Enshaeifar, *Student Member, IEEE*, Samaneh Kouchaki, *Student Member, IEEE*,
Clive Cheong Took, *Senior Member, IEEE*, and Saeid Sanei, *Senior Member, IEEE*

Abstract—A novel quaternion-valued singular spectrum analysis (SSA) is introduced for multichannel analysis of electroencephalogram (EEG). The analysis of EEG typically requires the decomposition of data channels into meaningful components despite the notoriously noisy nature of EEG - which is the aim of SSA. However, the singular value decomposition involved in SSA implies the strict orthogonality of the decomposed components, which may not reflect accurately the sources which exhibit similar neural activities. To allow for the modelling of such co-channel coupling, the quaternion domain is considered for the first time to formulate the SSA using the augmented statistics. As an application, we demonstrate how the augmented quaternion-valued SSA (AQSSA) can be used to extract the sources, even at a signal-to-noise ratio as low as -10 dB. To illustrate the usefulness of our quaternion-valued SSA in a rehabilitation setting, we employ the proposed SSA for sleep analysis to extract statistical descriptors for five-stage classification (Awake, N1, N2, N3 and REM). The level of agreement using these descriptors was 74% as quantified by the Cohen's kappa.

Index Terms—Electroencephalogram, singular spectrum analysis, quaternion, source extraction, sleep analysis.

I. INTRODUCTION

IN nature, observed signals are often mixtures of several sources corrupted by background noise. A common problem in neural system analysis, is therefore to recover the underlying components before any neurophysiologically meaningful information can be deduced from the electroencephalogram (EEG). Examples include extraction of P300 components in event related potential analysis, and estimation of brain rhythms for sleep analysis [1].

In the context of biomedical applications, singular spectrum analysis (SSA) has recently been introduced as a powerful model-free method for analysing real-valued data. SSA decomposes a signal into a number of interpretable components with different subspaces, such as trend, periodic, quasi-periodic and noise components [2], [3]. Unlike principal component analysis (PCA) where the number of subspaces is limited, the subspaces in SSA can be adjusted by varying the so-called parameter 'embedding dimension'.

Recently, SSA has been considered to solve several biomedical problems such as separation of EMG and ECG [4] and restoration of lung sound from heart sound [5]. In terms of rehabilitation applications, SSA has been used to remove

background biological processes so that the analysis of the movement responses in functional near-infrared spectroscopy (fNIRS) can be measured more accurately [6]. The challenging task in that particular application was to exploit the spectral overlap between the background activities and the movement related sources, illustrating the key advantage of SSA over traditional time-frequency and filtering methods. Another rehabilitation application of SSA has been the assessment of movement disabilities; SSA provided a way to measure the quality of a patient's upper limb movements [7]. More recently, the enhanced estimation of the brain rhythm alpha and delta by SSA enabled the accurate detection of two stages of sleep [8]. These applications illustrate the usefulness of SSA in biomedical and rehabilitation applications.

To date, SSA-based source extraction techniques are restricted to real-valued processing. However, the advances in multidimensional sensor technologies have highlighted the need for signal processing algorithms in complex and quaternion domain due to their great potential for modelling of two-dimensional (2-D) and 4-D data [9]. For instance, the complex-valued SSA enabled the detection of event related potential (ERP) components for the classification of schizophrenic patients [10] and improved forecasting of bivariate signals [11]. In the same spirit, the established common spatial pattern was recently introduced in the complex domain to model the correlation in EEG processing to classify motor imagery tasks [12]. As it is known that the constituent narrow-band EEG components originate from different brain zones [13], [14], two-channel processing offered by complex-valued techniques will be inadequate to model the correlation between those multiple processes. Instead, we consider the quaternion domain which generalises the complex domain to four dimensions to account for the multi-channel correlation.

To this end, we propose to exploit the latest advances in quaternion statistics to access the second order co-channels information [15] and quaternion singular value decomposition (SVD) to perform multi-dimensional decomposition [16]. This will allow us to formulate the quaternion-valued SSA in a rigorous way. Indeed, modelling based on quaternion augmented statistics have been successfully deployed in real-world applications such as wind forecasting [17] and extraction of ocular artefacts [9] and other applications [18]. We introduce the augmented quaternion SSA (AQSSA) and illustrate how it can solve the under-determined source extraction problem even when the number of sources exceeds the number of

electrodes. Moreover, it performs well even at a signal-to-noise ratio (SNR) less than -5 dB.

To evaluate the performance of the proposed method for EEG analysis, motivated by the success of our previous method [8], we consider sleep analysis as the application of interest. In the context of sleep analysis, SSA has recently been used to detect the transition between the two stages of sleep [8]. However, in contrast to [8], we consider five-stage classification. The basis for sleep EEG analysis lies in the estimation of the brain rhythms such as delta (less than 4 Hz), theta (4-7 Hz), alpha (8-12 Hz), **sigma** (11-16 Hz), and beta (16-30 Hz) from the EEG, that exhibit different characteristics in various stages of sleep, especially in the context of non-rapid eye movement (NREM) [1], [19]. NREM **sleep** is divided into three stages, namely N1, N2 and N3. Stage N1 is the drowsiness stage that is considered as a transition stage from awake (W) to asleep. In this stage, the alpha rhythm is attenuated and replaced by low amplitude waves, predominantly 4-7 Hz or theta, for more than 50% of the **epochs**. Next stage or Stage N2 is characterised by the occurrence of sleep spindles¹ or K complexes. Stage N2 might lead to Stage N3 which consists of slow wave activities (0.5-2 Hz or **lower** delta) measured frontally in adults [20]. The spindles may exist in this stage, however, as the sleep deepens the spindles' rate decreases [1]. Using these characteristics, we illustrate that quaternion SSA can be used to estimate five brain rhythms to track the NREM and REM **states** of sleep.

The paper is structured as follows: in Section II and III, a brief background on SSA algorithm and statistics of the quaternion domain are explained. Then, augmented quaternion SSA is introduced in Section IV, followed by the proposed supervised reconstruction methods. The results of applying the proposed methods to both synthetic and real data are shown in Section VI-VII and finally, Section VIII concludes the paper.

II. SINGULAR SPECTRUM ANALYSIS

SSA is a subspace decomposition algorithm that represents the original signal as a sum of components which can be identified as a trend, periodic or quasi-periodic component and noise. To achieve this, SSA algorithm is computed in two main stages which are described as follows [3].

A. Decomposition

The decomposition stage involves an embedding procedure followed by SVD of the covariance of the embedded data. For this purpose, the original signal $\mathbf{x} = [x_1, \dots, x_N] \in \mathbb{R}^{1 \times N}$ is mapped to the trajectory matrix $\mathbf{W} \in \mathbb{R}^{L \times K}$ where its first column is a segment of the original signal and its second column is the one-step lagged version of the first column; see Equation (1). The parameter L is defined as the window length where $1 < L < N$ and $K = N - L + 1$ where N is length of the original signal. L should be adequately long to account for the information about data variation. Observe that \mathbf{W} is a

Hankel matrix in which the elements along the cross-diagonals are equal.

$$\mathbf{x} = [x_1, \dots, x_N]$$

$$\mathbf{W} = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_K \\ x_2 & x_3 & x_4 & \dots & x_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & x_{L+2} & \dots & x_N \end{bmatrix} \in \mathbb{R}^{L \times K} \quad (1)$$

The trajectory matrix \mathbf{W} is then represented in terms of its eigenvectors and singular values which are obtained from the SVD of the covariance matrix $E\{\mathbf{W}\mathbf{W}^T\} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$, such that

$$\mathbf{W} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{V}^T$$

$$\mathbf{W} = \sum_{j=1}^r \mathbf{W}_j = \sum_{j=1}^r \sqrt{\lambda_j} \mathbf{u}_j \mathbf{v}_j^T \quad (2)$$

Observe that $\sqrt{\lambda_j}$ are the singular values of matrix \mathbf{W} in decreasing order ($\lambda_1 > \lambda_2 > \dots > \lambda_L$) and \mathbf{u}_j and \mathbf{v}_j are the left and right j -th eigenvectors. The parameter $r = \max\{j : \lambda_j > 0\}$ restricts the eigenvalues corresponding to the dominant components. The set $(\lambda_j, \mathbf{u}_j, \mathbf{v}_j)$ is called j -th eigentriple of matrix \mathbf{W} and \mathbf{W}_j is defined as the elementary matrix, i.e. $\mathbf{W}_j = \sqrt{\lambda_j} \mathbf{u}_j \mathbf{v}_j^T$.

B. Reconstruction

This stage reconstructs the signal from the elementary matrices, while mitigating the effect of noise. In the first part, the elementary matrices \mathbf{W}_j are divided into a number of groups and each group g is represented by a corresponding matrix $\hat{\mathbf{W}}_g$ which is the sum of elementary matrices within subset S_g :

$$\hat{\mathbf{W}}_g = \sum_{j \in S_g} \mathbf{W}_j = \begin{bmatrix} \hat{w}_{11} & \hat{w}_{12} & \dots & \hat{w}_{1,K} \\ \hat{w}_{21} & \hat{w}_{22} & \dots & \hat{w}_{2,K+1} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{w}_{L,1} & \hat{w}_{L,L+1} & \dots & \hat{w}_{L,N} \end{bmatrix} \quad (3)$$

Grouping of the elementary matrices is based on prior knowledge about the signal of interest. For example, for smoothing of the data, \mathbf{W}_j are divided into two groups which represent the dominant and noise subspaces; the noise subspace is ignored for reconstruction. Another example is when extracting the main trend and periodic components of the signal, the selected eigentriples correspond to the highest eigenvalue and paired eigenvalues respectively. Thus, in the second step of reconstruction a desired subset is selected and subseries $\hat{\mathbf{x}}$ is achieved by Hankelization algorithm, which computes the average along cross-diagonals of matrix $\hat{\mathbf{W}}_g$. In other words, if \hat{w}_{ij} refers to an entry of the matrix $\hat{\mathbf{W}}_g$, n -th term of the reconstructed signal \hat{x}_n is calculated by averaging \hat{w}_{ij} along all i, j such that $i + j = n + 1$ [3]:

$$\hat{\mathbf{x}} = [\hat{x}_1, \hat{x}_2, \dots, \hat{x}_N] \quad \text{where} \quad \begin{cases} \hat{x}_1 &= \hat{w}_{11} \\ \hat{x}_2 &= (\hat{w}_{12} + \hat{w}_{21})/2 \\ \hat{x}_3 &= (\hat{w}_{13} + \hat{w}_{22} + \hat{w}_{31})/3 \\ \dots & \end{cases} \quad (4)$$

¹Sleep spindles are oscillations in the sigma range (11-16 Hz) which last for short duration 0.5-2 seconds.

III. QUATERNION

A. Quaternion algebra

A quaternion vector \mathbf{x} comprises of a real (scalar) part $\Re\{\cdot\}$ and vector part (also called pure quaternion) $\Im\{\cdot\}$ which consists of three imaginary components:

$$\begin{aligned}\mathbf{x} &= \Re\{\mathbf{x}\} + \Im\{\mathbf{x}\} \\ &= \Re\{\mathbf{x}\} + i\Im_i\{\mathbf{x}\} + j\Im_j\{\mathbf{x}\} + k\Im_k\{\mathbf{x}\} \\ &= \mathbf{x}_a + i\mathbf{x}_b + j\mathbf{x}_c + k\mathbf{x}_d\end{aligned}\quad (5)$$

Note that the imaginary units i , j and k are orthogonal unit vectors for which:

$$\begin{aligned}ij &= -ji = k & jk &= -kj = i & ki &= -ki = j \\ i^2 &= j^2 = k^2 = ijk = -1\end{aligned}\quad (6)$$

The above identities demonstrate the noncommutative property of quaternion products where $\mathbf{x}_1\mathbf{x}_2 \neq \mathbf{x}_2\mathbf{x}_1$. Another important notion for the quaternion domain (\mathbb{H}) is the so-called ‘‘quaternion involution’’ which defines a self-inverse mapping. All three involutions about the i , j and k imaginary unit axes are given by [15]:

$$\begin{aligned}\mathbf{x}^i &= -i\mathbf{x}i = \mathbf{x}_a + i\mathbf{x}_b - j\mathbf{x}_c - k\mathbf{x}_d \\ \mathbf{x}^j &= -j\mathbf{x}j = \mathbf{x}_a - i\mathbf{x}_b + j\mathbf{x}_c - k\mathbf{x}_d \\ \mathbf{x}^k &= -k\mathbf{x}k = \mathbf{x}_a - i\mathbf{x}_b - j\mathbf{x}_c + k\mathbf{x}_d\end{aligned}\quad (7)$$

which form the bases for augmented quaternion statistics. Note that involution illustrates a rotation along a single unit axis, while the quaternion conjugate operator $(\cdot)^*$ rotates along all three imaginary axes:

$$\mathbf{x}^* = \Re\{\mathbf{x}\} - \Im\{\mathbf{x}\} = \mathbf{x}_a - i\mathbf{x}_b - j\mathbf{x}_c - k\mathbf{x}_d \quad (8)$$

The relationship between a quaternion vector and its involutions is given by:

$$\mathbf{x}^* = \frac{1}{2}[\mathbf{x}^i + \mathbf{x}^j + \mathbf{x}^k - \mathbf{x}] \quad (9)$$

Observe that as long as four of the five vectors in (9) are known, the fifth vector can be computed.

B. Augmented statistics

To exploit the power difference or the correlation between the data channels, it is necessary to make use of widely linear modelling of signals. For this purpose ‘‘augmented’’ statistics have been established to incorporate the complementary covariance matrices and exploit the complete second-order information.

In complex domain, real and imaginary parts of a complex number $\mathbf{z} = \mathbf{z}_a + i\mathbf{z}_b$ can be calculated as $\mathbf{z}_a = \frac{1}{2}(\mathbf{z} + \mathbf{z}^*)$ and $\mathbf{z}_b = \frac{1}{2i}(\mathbf{z} - \mathbf{z}^*)$. These two examples highlight the necessity to consider both \mathbf{z} and \mathbf{z}^* , paving the way for augmented complex statistics, where a basis vector is *augmented* by its conjugate variables as $\mathbf{z}^a = [\mathbf{z}, \mathbf{z}^*]^T \in \mathbb{C}$ [15]. Thus, the augmented covariance matrix \mathbf{C}^a can be computed as:

$$\mathbf{C}^a = \mathbf{z}^a \mathbf{z}^{aH} = \begin{bmatrix} \mathbf{z}\mathbf{z}^H & \mathbf{z}\mathbf{z}^T \\ (\mathbf{z}\mathbf{z}^T)^* & (\mathbf{z}\mathbf{z}^H)^* \end{bmatrix} \quad (10)$$

Note that the Hermitian operator $(\cdot)^H$ refers to transpose conjugate operation. Since \mathbf{C}^a includes both covariance and

pseudo-covariance, it takes into account the complete second-order information in the complex-valued data. However, in quaternion domain such convenient manipulation is not possible. The correspondence between the elements of a quaternion variable in \mathbb{H} and the elements of a quadrivariate vector in \mathbb{R}^4 can be obtained by [15]:

$$\begin{aligned}\mathbf{x}_a &= \frac{1}{2}(\mathbf{x} + \mathbf{x}^*) & \mathbf{x}_b &= \frac{1}{2i}(\mathbf{x} - \mathbf{x}^{i*}) \\ \mathbf{x}_c &= \frac{1}{2j}(\mathbf{x} - \mathbf{x}^{j*}) & \mathbf{x}_d &= \frac{1}{2k}(\mathbf{x} - \mathbf{x}^{k*})\end{aligned}\quad (11)$$

Thus, the quaternion statistics should include all quaternion involutions \mathbf{x}^i , \mathbf{x}^j and \mathbf{x}^k to access to the complete second order statistical information. This way, the augmented quaternion vector $\mathbf{x}^a = [\mathbf{x}^T, \mathbf{x}^{iT}, \mathbf{x}^{jT}, \mathbf{x}^{kT}]^T$ is defined and used to compute the augmented covariance matrix (\mathbf{C}_q^a):

$$\begin{aligned}\mathbf{C}_q^a &= E\{\mathbf{x}^a \mathbf{x}^{aH}\} \\ &= \begin{bmatrix} \mathbf{C}_{\mathbf{xx}} & \mathbf{C}_{\mathbf{x}^i} & \mathbf{C}_{\mathbf{x}^j} & \mathbf{C}_{\mathbf{x}^k} \\ \mathbf{C}_{\mathbf{x}^i}^H & \mathbf{C}_{\mathbf{x}^i \mathbf{x}^i} & \mathbf{C}_{\mathbf{x}^i \mathbf{x}^j} & \mathbf{C}_{\mathbf{x}^i \mathbf{x}^k} \\ \mathbf{C}_{\mathbf{x}^j}^H & \mathbf{C}_{\mathbf{x}^j \mathbf{x}^i} & \mathbf{C}_{\mathbf{x}^j \mathbf{x}^j} & \mathbf{C}_{\mathbf{x}^j \mathbf{x}^k} \\ \mathbf{C}_{\mathbf{x}^k}^H & \mathbf{C}_{\mathbf{x}^k \mathbf{x}^i} & \mathbf{C}_{\mathbf{x}^k \mathbf{x}^j} & \mathbf{C}_{\mathbf{x}^k \mathbf{x}^k} \end{bmatrix}\end{aligned}\quad (12)$$

in which $\mathbf{C}_{\mathbf{xx}} = E\{\mathbf{xx}^H\}$ is the standard covariance matrix, whereas the matrices $\mathbf{C}_{\mathbf{x}^i}$, $\mathbf{C}_{\mathbf{x}^j}$, and $\mathbf{C}_{\mathbf{x}^k}$ are respectively termed i -, j -, and k - covariance matrices. These matrices are also called the complementary or pseudo-covariance matrices [15]:

$$\begin{aligned}\mathbf{C}_{\mathbf{x}^\alpha} &= E\{\mathbf{xx}^{\alpha H}\} & \mathbf{C}_{\mathbf{x}^\alpha \mathbf{x}^\beta} &= E\{\mathbf{x}^\alpha \mathbf{x}^{\beta H}\} \\ & \text{where } \alpha, \beta \in \{i, j, k\}\end{aligned}\quad (13)$$

As \mathbf{C}_q^a considers both covariance and pseudo-covariance matrices, it takes into account the complete second order information in the quaternion-valued data.

IV. AUGMENTED QUATERNION SSA (AQSSA)

The basic quaternion trajectory matrix $\mathbf{W} \in \mathbb{H}^{L \times K}$ is generated via (1). The next step aims at calculating quaternion SVD (QSVD) of the trajectory matrix using its covariance matrix. Unlike real-valued SSA in which the covariance was generated as $E\{\mathbf{W}\mathbf{W}^T\}$, for AQSSA the basic trajectory matrix incorporates information augmented by all three involutions (7) to generate the augmented trajectory matrix $\mathbf{W}^a \in \mathbb{H}^{4L \times 4K}$ as:

$$\mathbf{W}^a = [\mathbf{W}^T, \mathbf{W}^{iT}, \mathbf{W}^{jT}, \mathbf{W}^{kT}]^T \quad (14)$$

where $\mathbf{W}^{\alpha T}$, $\alpha \in \{i, j, k\}$, is transpose of the α -involution operation of the trajectory matrix \mathbf{W} . The generated \mathbf{W}^a is then used to compute the new augmented covariance matrix $\mathbf{C}^a \in \mathbb{H}^{4L \times 4L}$.

$$\begin{aligned}\mathbf{C}^a &= E\{\mathbf{W}^a \mathbf{W}^{aH}\} \\ &= \begin{bmatrix} \mathbf{C}_{\mathbf{W}\mathbf{W}} & \mathbf{C}_{\mathbf{W}^i} & \mathbf{C}_{\mathbf{W}^j} & \mathbf{C}_{\mathbf{W}^k} \\ \mathbf{C}_{\mathbf{W}^i}^H & \mathbf{C}_{\mathbf{W}^i \mathbf{W}^i} & \mathbf{C}_{\mathbf{W}^i \mathbf{W}^j} & \mathbf{C}_{\mathbf{W}^i \mathbf{W}^k} \\ \mathbf{C}_{\mathbf{W}^j}^H & \mathbf{C}_{\mathbf{W}^j \mathbf{W}^i} & \mathbf{C}_{\mathbf{W}^j \mathbf{W}^j} & \mathbf{C}_{\mathbf{W}^j \mathbf{W}^k} \\ \mathbf{C}_{\mathbf{W}^k}^H & \mathbf{C}_{\mathbf{W}^k \mathbf{W}^i} & \mathbf{C}_{\mathbf{W}^k \mathbf{W}^j} & \mathbf{C}_{\mathbf{W}^k \mathbf{W}^k} \end{bmatrix}\end{aligned}\quad (15)$$

where $\mathbf{C}_{\mathbf{W}\mathbf{W}}$ is the standard covariance matrix and the matrices $\mathbf{C}_{\mathbf{W}^i}$, $\mathbf{C}_{\mathbf{W}^j}$, and $\mathbf{C}_{\mathbf{W}^k}$ are the complementary matrices.

Algorithm 1 AQSSA algorithm

Decomposition

$$\mathbf{x} = [x_1, \dots, x_N] \in \mathbb{H}$$

$$\mathbf{W} = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_K \\ x_2 & x_3 & x_4 & \dots & x_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & x_{L+2} & \dots & x_N \end{bmatrix} \in \mathbb{H}^{L \times K}$$

$$\mathbf{W}^a = [\mathbf{W}^T, \mathbf{W}^{iT}, \mathbf{W}^{jT}, \mathbf{W}^{kT}]^T \in \mathbb{H}^{4L \times K}$$

$$\mathbf{C}^a = E\{\mathbf{W}^a \mathbf{W}^{aH}\}$$

$$= \begin{bmatrix} \mathbf{C}_{\mathbf{W}\mathbf{W}} & \mathbf{C}_{\mathbf{W}^i} & \mathbf{C}_{\mathbf{W}^j} & \mathbf{C}_{\mathbf{W}^k} \\ \mathbf{C}_{\mathbf{W}^i}^H & \mathbf{C}_{\mathbf{W}^i \mathbf{W}^i} & \mathbf{C}_{\mathbf{W}^i \mathbf{W}^j} & \mathbf{C}_{\mathbf{W}^i \mathbf{W}^k} \\ \mathbf{C}_{\mathbf{W}^j}^H & \mathbf{C}_{\mathbf{W}^j \mathbf{W}^i} & \mathbf{C}_{\mathbf{W}^j \mathbf{W}^j} & \mathbf{C}_{\mathbf{W}^j \mathbf{W}^k} \\ \mathbf{C}_{\mathbf{W}^k}^H & \mathbf{C}_{\mathbf{W}^k \mathbf{W}^i} & \mathbf{C}_{\mathbf{W}^k \mathbf{W}^j} & \mathbf{C}_{\mathbf{W}^k \mathbf{W}^k} \end{bmatrix}$$

↓ Quaternion SVD

$$\mathbf{W}^a = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{V}^H$$

$$\mathbf{W}^a = \sum_{j=1}^r \mathbf{W}_j^a = \sum_{j=1}^r \sqrt{\lambda_j} \mathbf{u}_j \mathbf{v}_j^H$$

where $r = \max\{j : \lambda_j > 0\}$

Reconstruction

$$\hat{\mathbf{W}}_g^a = \sum_{j \in S_g} \mathbf{W}_j^a = \begin{bmatrix} \hat{w}_{11} & \hat{w}_{12} & \dots & \hat{w}_{1,K} \\ \hat{w}_{21} & \hat{w}_{22} & \dots & \hat{w}_{2,K+1} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{w}_{L,1} & \hat{w}_{L,L+1} & \dots & \hat{w}_{L,N} \end{bmatrix}$$

$$\hat{\mathbf{x}} = [\hat{x}_1, \hat{x}_2, \dots, \hat{x}_N] \quad \text{where} \quad \begin{cases} \hat{x}_1 = \hat{w}_{11} \\ \hat{x}_2 = (\hat{w}_{12} + \hat{w}_{21})/2 \\ \hat{x}_3 = (\hat{w}_{13} + \hat{w}_{22} + \hat{w}_{31})/3 \\ \dots \end{cases}$$

Thus, \mathbf{C}^a captures the complete second order information and it is more likely that AQSSA outperforms its basic quaternion-valued SSA (QSSA). The steps of AQSSA are summarised in Algorithm 1.

V. SUPERVISED RECONSTRUCTION

To this point, AQSSA has been used to decompose the four correlated channels simultaneously into several subspaces. In order to employ this method for source extraction, the corresponding subspace of the desired signal should be identified. Similar to the real-valued SSA, one of the main challenges is to find S_g , the group of the eigentriples for reconstructing the component of interest. Thus, it is necessary to develop some techniques for proper subspace grouping [8].

Subspaces are generally characterised by various statistical or physical constraints based on desired properties. In this work we aim to develop an under-determined source extraction method assuming that the sources are narrow-band and their frequencies are within a limited range. To achieve this, two possible approaches are discussed in the following sections:

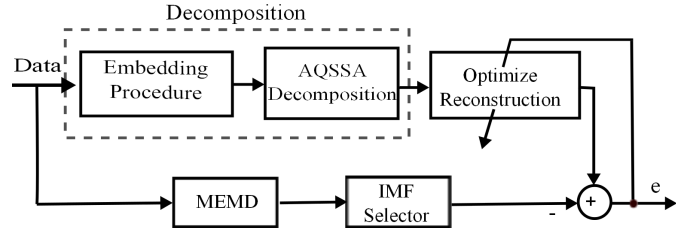


Fig. 1: Supervised AQSSA using an adaptive procedure for selection of the desired subspace with the help of MEMD.

A. Generalisation of our previous work to multi channels

Recently, empirical mode decomposition (EMD) has been used in the reconstruction stage of a tensor-based SSA (TSSA) to accurately select the desired subspace when frequency band of the signal of interest is known [8]. The EMD tensor-based SSA (TSSA-EMD) has been introduced for single channel source separation with the application in sleep scoring [8]. In this work, however, the multivariate EMD (MEMD) [21] is used in the reconstruction stage (3) to inform AQSSA on the optimal grouping subset S_g .

Similar to traditional EMD, MEMD is a time-frequency analysis tool in which the original multivariate signal is represented as a linear combination of intrinsic oscillatory modes, called intrinsic mode functions (IMFs). Since MEMD involves multivariate data, the signal is first projected in different directions to generate multiple envelopes. These envelopes are then averaged to calculate the local mean to yield the IMFs [21]. Since MEMD estimates the number of frequency components within each subspace, it is used to select the subset of the desired signal. For this purpose, a number of IMFs are selected based on some prior criteria, such as the maximum power in the desired frequency band. Then, the following equation is used to automatically group the eigentriples:

$$J = \|\mathbf{F} - \mathbf{U}\mathbf{Y}\mathbf{\Lambda}^{1/2}\mathbf{V}^H\| \quad (16)$$

in which \mathbf{F} is the Hankel matrix of the selected IMFs, \mathbf{Y} is a diagonal matrix of adaptive weights and J is the objective function. This function can be minimised using the gradient descent to obtain \mathbf{Y} :

$$\mathbf{Y}^{\text{new}} = \mathbf{Y}^{\text{old}} - \eta \nabla_{\mathbf{Y}^{\text{old}}}(J) \quad (17)$$

where η is the step size set empirically and $\nabla_{\mathbf{Y}^{\text{old}}}(J)$ denotes the gradient of the objective function J with respect to \mathbf{Y} [8]. The structure of AQSSA-MEMD is summarised in Fig. 1.

Remark#1: Unlike TSSA-EMD which is a single channel source separation method, the generalised AQSSA-MEMD method can cater for multi-channel processing, typically used in brain computer interface applications.

Remark#2: Note that AQSSA-MEMD can extract an arbitrary number of sources from four recorded channels as long as the number of sources is less than L .

Remark#3: TSSA-EMD can track the changes of delta and alpha band in sleep analysis [8]. However, AQSSA-MEMD can estimate five frequency ranges of brain rhythms. Thus, the proposed method can provide a more refined extraction of delta, theta, alpha, **sigma**, and beta bands from the EEGs.

B. A novel multi-channel algorithm

In addition to the novel quaternion algorithm presented in Section V-A, we introduce a second algorithm that circumvents the optimisation step (17), which requires selection of the optimal step size. Indeed, selection of the optimal step size generally involves a lengthy trial and error procedure. To address this shortcoming, we propose a fully data-driven technique called single-element AQSSA (AQSSA-SE) in which the reconstruction stage is supervised using a frequency constraint. Assume that $\mathbf{x} \in \mathbb{H}$ contains four data channels which are obtained from several sources using a random mixing matrix \mathbf{A} :

$$\begin{aligned} \mathbf{x} &= \mathbf{x}_a + i\mathbf{x}_b + j\mathbf{x}_c + k\mathbf{x}_d \\ \begin{bmatrix} \mathbf{x}_a \\ \mathbf{x}_b \\ \mathbf{x}_c \\ \mathbf{x}_d \end{bmatrix} &= \mathbf{A}\mathbf{s} \quad \text{where} \quad \begin{cases} \mathbf{A} \in \mathbb{R}^{4 \times D} \\ \mathbf{s} \in \mathbb{R}^{D \times N} \end{cases} \end{aligned} \quad (18)$$

where D defines the number of sources. The objective is to implement AQSSA-SE to reconstruct the sources $\hat{\mathbf{s}}$ from \mathbf{x} . To this end, the original signal is first represented in the eigen domain with $4L$ eigentriple sets, see Algorithm 2. In the reconstruction procedure, the eigen domain is divided into two subspaces, i.e. $[1, r]$ and $(r, 4L]$, termed as dominant (lower) and noise (higher) subspaces respectively. To mitigate the effect of noise, the latter subspace is ignored and grouping is only applied to the dominant subspace.

For an automatic grouping, a single set of eigentriples is used at a time to reconstruct a quaternion-valued signal $\hat{\mathbf{x}}$. Since the sources are considered narrow-band with known frequency ranges, the corresponding eigenvector can be automatically placed in a proper group according to its dominant frequency obtained from the Fourier transform. These steps are repeated for all the eigentriples in the lower subspace. Note that there are a total number of D groups, which is the same number as the original sources and each group contains an arbitrary number of components (see Algorithm 2, Reconstruction step). These components are then used to extract a unique source by computing

$$\hat{\mathbf{s}}_d = \sum_{i=1}^p \text{sign}(\text{corr}(\hat{\mathbf{x}}_m, \hat{\mathbf{x}}_i)) \hat{\mathbf{x}}_i \quad (19)$$

in which the function ‘sign’ represents sign of its input and ‘corr’ measures the correlation between the corresponding inputs. Note that $d = [1, \dots, D]$ is the group index, p is the total number of components in each group and $\hat{\mathbf{x}}_m$ has the maximum power in the corresponding group d . The steps of AQSSA-SE are summarised in Algorithm 2.

Remark#4: Similar to AQSSA-MEMD introduced in Section V-A, the AQSSA-SE method can be applied to 4D data to extract arbitrary number of sources. Thus, it can address an under-determined source separation scenario.

Remark#5: Since the proposed AQSSA-SE method considers the correlation between the data, the accurate ratio of the sources in different channels is preserved in the unmixing matrix. However, it suffers from the scaling problem - the result of extracting and adding the effect of desired source in all the channels using Equation (19).

Algorithm 2 AQSSA-SE algorithm for under-determined BSS

Decomposition

$\mathbf{x} \rightarrow \mathbf{W} \rightarrow \mathbf{W}^a \in \mathbb{H}^{4L \times K} \rightarrow E\{\mathbf{W}^a \mathbf{W}^{aH}\} \rightarrow (\lambda_j, \mathbf{u}_j, \mathbf{v}_j)$
The decomposition stage follows that of the AQSSA technique. For a full mathematical description, see Algorithm 1.

Single-element reconstruction

1. Select $r = \frac{L}{2}$ to highlight the dominant subspace $[1, r]$.
2. Generate a single elementary matrix $\hat{\mathbf{W}}^a \approx \sqrt{\lambda_j} \mathbf{u}_j \mathbf{v}_j^H$ where $j = [1, \dots, r]$.
3. Reconstruct the signal $\hat{\mathbf{x}} \in \mathbb{H}$ by Hankelization algorithm, Equation (4).
4. Place $\hat{\mathbf{x}}$ in a proper group according to its frequency.
5. Repeat steps 2-4 for all the eigentriples in the dominant subspace $[1, r]$.

Source extraction

6. Select a component $\hat{\mathbf{x}}_m$ with maximum power in each group.
7. Extract the source: $\hat{\mathbf{s}}_d = \sum_{i=1}^p \text{sign}(\text{corr}(\hat{\mathbf{x}}_m, \hat{\mathbf{x}}_i)) \cdot \hat{\mathbf{x}}_i$
8. Repeat steps 6-7 for all groups $[1, \dots, D]$.

Remark#6: Furthermore, AQSSA-SE filters out noise and performs source extraction, while AQSSA-MEMD only extracts the sources. It is thus expected that AQSSA-SE provides better estimation of noisy sources.

VI. EXPERIMENTAL RESULTS

In this section, performances of the proposed methods are evaluated using simulated data. First part explains the parameter setting and the evaluation criteria; then, the proposed methods are applied to synthetic data. In order to highlight the advantage of augmented quaternion statistics, AQSSA-SE and AQSSA-MEMD were compared with QSSA-SE and QSSA-MEMD.

A. Simulations

The first experiment was carried out to evaluate the reconstruction accuracy for different levels of white Gaussian noise. To generate the synthetic EEG data, we followed the example of Koles and Soong [22] in which sum of sinusoids were used to simulate EEG to validate their source localization. For this purpose, five narrow-band signals (with frequency characteristics of delta, theta, alpha, **sigma**, and beta) were generated with a length of $N = 1000$ samples. Then, they were mixed using a 4×5 random matrix drawn from the standard normal distribution to obtain four mixtures. Overall, 20 sets of simulated data were generated to compute the averaged results. Furthermore, in order to assess the effect of noise on the performance accuracy, white Gaussian noise with different levels of SNR varying from -10 dB to 10 dB was added to each set. The simulated signals were modelled as:

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{b}\mathbf{e} \quad (20)$$

where \mathbf{e} is the white Gaussian noise, b determines the noise level, \mathbf{x} is the generated input and \mathbf{s} is the desired sources to be extracted. In this work, four observed channels were mixture of five sources. These four channels were represented as a single quaternion channel to generate the input of our experiment.

The performance of proposed methods were evaluated by considering normalised RMS error (NRMSE) among the initial source \mathbf{s}_d and the corresponding extracted one $\hat{\mathbf{s}}_d$. The performance measurement NRMSE was evaluated with regards to the noise free signal as:

$$\text{NRMSE} = \frac{\sqrt{\frac{\sum(\mathbf{s}_d - \hat{\mathbf{s}}_d)^2}{N}}}{\max(\mathbf{s}_d) - \min(\mathbf{s}_d)} \quad (21)$$

where $\hat{\mathbf{s}}_d$ indicates the d -th estimated source and N is the data length. Hence, the lower the error, the better the estimate. Moreover, the Pearson correlation level between the desired and the extracted sources was considered as another criterion to assess the performance:

$$\rho = \text{corr}(\mathbf{s}_d, \hat{\mathbf{s}}_d) = \frac{\sum(\mathbf{s}_d - \bar{\mathbf{s}}_d)(\hat{\mathbf{s}}_d - \bar{\hat{\mathbf{s}}}_d)}{\sqrt{\sum(\mathbf{s}_d - \bar{\mathbf{s}}_d)^2 \sum(\hat{\mathbf{s}}_d - \bar{\hat{\mathbf{s}}}_d)^2}} \quad (22)$$

in which ρ represents the correlation between the source and its estimate. Note that $\bar{\mathbf{s}}_d$ and $\bar{\hat{\mathbf{s}}}_d$ indicate respectively the mean values of d -th desired and estimated source.

1) *Parameter Settings:* The QSSA and AQSSA window length was set to $L = \frac{N}{4}$ in the experiments. The window length is related to 1/bandwidth of the data and it was ensured that L was not less than 2 or 3 times this value to cater for the information about data variation. However, very large values of L were avoided to have low computational complexity. In MEMD-based experiment, 8 to 13 IMFs were generated among which sum of IMFs within the frequency range of interest was used. In SE-based method, the dominant subspace was selected as $r = \frac{L}{2}$. Note that selecting a very small r would ignore the useful information, whereas very large values would increase the effect of noise subspace.

After selecting the appropriate parameters, QSSA-SE, QSSA-MEMD, AQSSA-SE, and AQSSA-MEMD were employed to extract the original sources. Figs. 2 and 3 show the NRMS error and correlation against SNR.

B. Simulation Discussion

The simulation results in Fig. 2 and 3 confirmed that the proposed method was able to extract the desired sources when the number of sources exceeded the number of electrodes. As expected, augmented quaternion methods outperformed the normal quaternion algorithms for separation of narrow-band data. The main reason is that augmented statistics accounted for the correlation between data channels by exploiting both covariance and pseudo-covariance matrices. Using the augmented statistics, less error was observed and there was higher degree of correlation between the extracted and desired sources as shown in Figs. 2 and 3.

Two reconstruction schemes were then considered. The results in Fig. 2 and 3 showed that the single-element reconstruction approach outperformed the MEMD-based methods. It

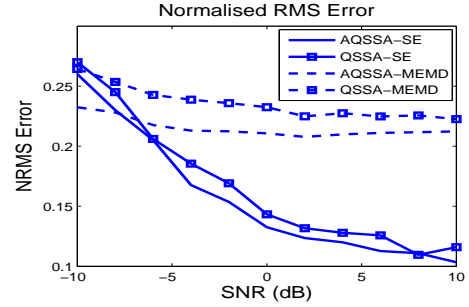


Fig. 2: The effect of noise level on the performance of QSSA-SE, QSSA-MEMD, AQSSA-SE, and AQSSA-MEMD for extraction of the simulated narrow-band signals. The error is averaged over 20 sets of data.

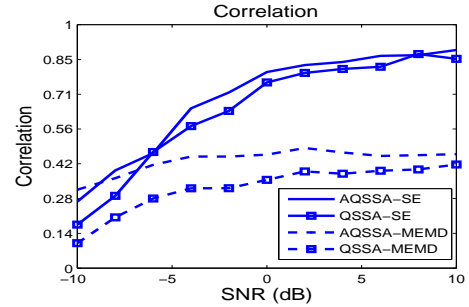


Fig. 3: The effect of noise level on the correlation between the desired and the estimated source signals averaged over 20 sets of simulated narrow-band data using QSSA-SE, QSSA-MEMD, AQSSA-SE, and AQSSA-MEMD.

is due to the fact that in the SE-based reconstruction, the noise subspace was removed and the outputs were more similar to the original sources. On the other hand, all the subspaces were considered simultaneously by MEMD. Unlike the SE-based method, MEMD had mode mixing, i.e. similar oscillations in nearby modes. Thus, it was reasonable to expect better performance from AQSSA-SE and this was also confirmed by the simulation results shown in Fig. 4. This figure illustrates the high performance of AQSSA-SE for extraction of sources from the noisy mixtures shown in Fig. 4 (b). Note that the estimated sources had approximately 80% correlation with the original sources, Fig. 4 (a),(c).

VII. APPLICATION TO SLEEP EEG

This section involves AQSSA-SE with application to sleep data. Sleep is a dynamic process during which little response to external stimuli and body movement and alteration in metabolic rate can be observed [1]. Sleep has two distinct states with various neural activity levels; non-rapid eye movement (NREM) and rapid eye movement (REM) sleep. Each **state** has different neurological and physiological characteristics. Sleep analysis and classification of sleep stages are specifically important for clinical diagnosis of sleep-related phenomena [20]. Furthermore, sleep analysis has been used for automatic detection of drowsiness during driving [23]. To identify the sleep stages, polysomnography (PSG) signals including EEG, EOG and EMG are recorded and labelled by a sleep expert according to the standard rules. The most common sleep scoring guideline is the American Academy of Sleep

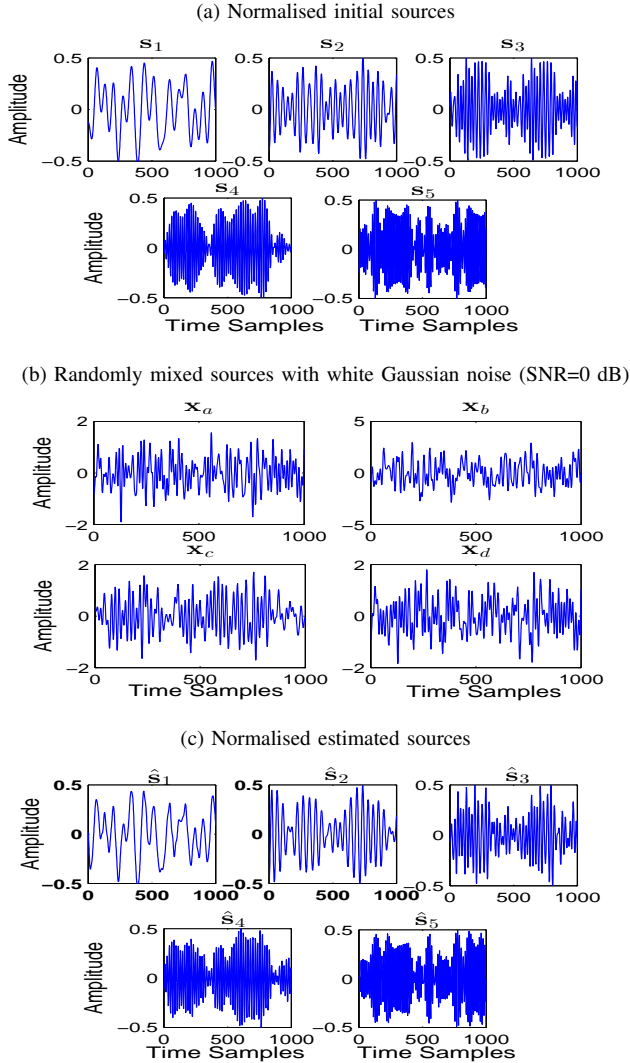


Fig. 4: Extracting five sources from four channels; (a) narrow-band desired sources, (b) random mixtures with added white Gaussian noise, and (c) the estimated sources using AQSSA-SE.

Medicine (AASM) manual which is designed to provide a comprehensive standardised specification and scoring rules for characterising natural sleep [20], [23].

As manual sleep scoring is a time-consuming process, several techniques have been established to automate sleep scoring [24], [25], [26]. In this paper, we aim to decompose the sleep EEG data and classify the stages using a multiclass support vector machine (M-SVM). SVM is a supervised learning method for classification which simultaneously maximise the classification margin and minimise the classification error [26]. Note that the main objective of this work is to introduce a novel multichannel analysis algorithm to exploit the correlation within bio-data, typically used for rehabilitation purposes. Thus, sleep analysis is performed to illustrate the applicability of our proposed method in EEG processing.

The data used in this work is a subset of the EEG data recorded in the Sleep Centre of the University of Surrey¹. The

¹Data is recorded and validated under supervision of clinical experts. The same dataset has been used in our previous work [8].

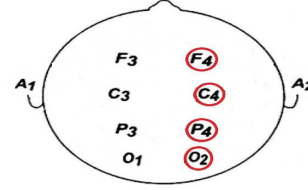


Fig. 5: The location of available electrodes. The electrodes used in this work are indicated by circles. Note that the selected channels are all referenced with respect to A1.

dataset recorded from thirty-six healthy men and women, all of which participated in EEG recording for nine sessions. During each session, PSG measures were recorded at a sampling rate of 256 Hz for a baseline night (BL, 8 hours), seven condition nights, i.e. sleep extension (SE, 10 hours) and sleep restriction (SR, 6 hours), and a recovery night (12 hours) following a period of total sleep deprivation.

To perform sleep classification, the AQSSA-SE method and a wavelet-based approach, which is a common method for analysis of EEG subbands [27], [28], were employed on sleep EEG to extract frequency bands delta, theta, alpha, **sigma**, and beta. Morlet wavelet transform which was used for comparison in this work, is an established approach for providing high time-frequency resolution in sleep signal analysis especially for detection of sleep abnormalities and spindles [28]. As wavelet works with single channel data, it was applied on four channels (F4-A1, C4-A1, P4-A1, and O2-A1) individually and results were averaged to generate the final output. However, the proposed method considered these four channels simultaneously, see Fig. 5 for the location of these electrodes.

Remark#7: According to the AASM manual, the right-sided electrodes are recommended for sleep analysis. Furthermore, it was shown in [29] that delta, theta, and beta bands exhibit right lateralisation during the sleep. Thus, in this work right-sided electrodes were selected to capture all the necessary information.

Remark#8: The selected electrodes provided a reasonable coverage of the frontal to posterior brain regions. Thus, it was expected that this method showed clear outputs for all frequency range sources regardless of their locations.

Remark#9: A complete automatic sleep scoring system should consider full PSG measures including EEG, EOG and EMG [20], [24]. However, in this work only EEG channels were considered without the assistance of EMG or EOG information to focus on EEG processing.

A. Subband decomposition and statistical descriptors

EEG data often contains many events, movement related sources, artifacts and noise. Therefore, it is often required to solve a noisy source separation problem, for which we have proposed a quaternion-based method in under-determined source separation settings, described in Section V-B.

1) *Subband decomposition:* In the first step of sleep scoring, the EEG data were decomposed into specific frequency bands. For this purpose, AQSSA-SE and wavelet methods were applied over segments of two seconds (2-s) to decompose

TABLE I: Sleep stages description.

Stage	Description
Awake (W)	Alpha rhythm: 8-13 Hz activity over the occipital region with eye closure. In stage W with eye open, brain rhythms contain low-amplitude activity (mainly beta and alpha frequencies), Eye blinks: vertical eye movements at frequencies of 0.5-2 Hz, reading eye movements or voluntary rapid eye movements (REM), Stage W may range from full alertness through early stages of drowsiness.
Stage N1	Low amplitude, mixed frequency: predominantly 4-7 Hz activity, Slow eye movements (SEM), N1 is a transition stage between wakefulness and sleep, in which the alpha rhythm is attenuated and replaced by low amplitude mixed frequency waves.
Stage N2	Sleep spindle: a train of distinct waves at frequencies 11-16 Hz which lasts for short duration of 0.5-2 seconds. Spindles start around Stage N2 and its occurrence rate decreases from this stage to Stage N3, K complex: a negative sharp-wave followed by a positive component standing out from the background EEG.
Stage N3	Slow wave activity: waves of 0.5-2 Hz recorded over the frontal regions, Stage N3 replaces the <i>R&K</i> nomenclature of stage III and stage IV sleep, Note that sleep spindles may still exist in stage N3.
Stage R	Rapid eye movements: conjugate, irregular, sharply peaked eye movements, Low amplitude, mixed frequency EEG, Low chin EMG tone: during stage R, baseline EMG activity in the chin is usually at the lowest level of the entire recording, Transient muscle activity: irregular bursts of EMG activity which can be seen in the chin or anterior tibial EMG, as well as in EEG or EOG derivations.

each segment into five subbands delta, theta, alpha, **sigma**, and beta. In the next step, EEG signals were analysed independently within each band to extract spectral and temporal information [25], see Table I. Thus, the mean power and relative power of each subband were measured for each 2-s window.

2) *Statistical descriptors and classification:* To identify the sleep stages, a set of fifteen segments was considered as a 30-s epoch and the M-SVM classified the generated epochs. M-SVM is a learning method which constructs a model based on the training data and their corresponding labels. Thus, it was required to feed the M-SVM with specific features obtained from the training epochs [23], [26]. The statistical descriptors or features used in this work are selected according to the AASM manual [20] and the established automatic sleep scoring systems in [24], [25], [30]. **Each epoch was characterised by a set of 20 features which are summarised below:**

- Standard deviation of ‘mean power’ for frequency bands delta, theta, alpha and, beta over a group of fifteen 2-s segments.
- Standard deviation of ‘relative power’ for all subbands delta, theta, alpha, **sigma**, and beta over a group of fifteen 2-s segments.
- Density of ‘mean power’ for frequency bands delta, theta, alpha, and beta. The density of each band was calculated

TABLE II: Qualitative interpretation of κ .

Agreement	Kappa value	Agreement	Kappa value
Almost Perfect	>0.8	Fair	0.21-0.40
Substantial	0.61-0.8	Slight	0.00-0.20
Moderate	0.41-0.60	Poor	<0

by counting the number of 2-s segments with ‘normalised power’² above a specific threshold. The threshold was selected as 0.3 for delta and theta and 0.5 for alpha and beta. For instance, the density of alpha was calculated as the number of 2-s segments with normalised alpha power above 0.5.

- Density of ‘relative power’ for all subbands.
- Standard deviation and density of the power ratio $\frac{\text{beta} \times \text{theta}}{\text{delta}}$, for discrimination of REM from EEG as established in [25].

To estimate the sleep stages obtained from M-SVM classifier, two smoothing rules [25], [30] were adopted:

- Three consecutive rules to replace REM-N2-REM and REM-W-REM with REM-REM-REM and replace N2-N1-N2 with N2-N2-N2.
- Any REM epoch detected before the appearance of stage N2, was replaced with the wakefulness epoch.

B. Statistical comparison

Performance of the automated scoring using the AQSSA-SE method was evaluated using the following complementary measures which are the most common criteria used in the literature [25], [30].

Epoch-by-epoch agreement: Epoch-by-epoch agreement is defined as the percentage of epochs which are assigned the same stage in the manual and the automated scoring system.

Cohen’s kappa: Cohen’s kappa (κ) is a statistical measure which quantifies the inter-rater agreement between two or more scorers. As κ considers the probability of agreement due to chance alone, it is generally known to be more robust than simple percentage agreement [25]. The interpretation of κ is summarized in Table II.

Sensitivity: In addition to Cohen’s kappa and epoch-by-epoch agreement, the sensitivity associated with each stage was also calculated. Sensitivity is the **ratio of the number** of stages which were correctly detected by our proposed method³. Sensitivity reflects the performance compared to the current reference.

C. Results

To demonstrate the generalisation of the proposed method, the selected EEG channels (F4-A1, C4-A1, P4-A1, and O2-A1) for half of the subjects were used to train M-SVM and the other half of the data were used to test the reconstructed model.

²Normalisation was performed for an individual subband with respect to its maximum power within a specific epoch.

³For example, sensitivity of awake in Table III was calculated as $\frac{222}{222+45}$.

TABLE III: Confusion matrix for AQSSA-SE method

		AQSSA-SE					Sensitivity
		Awake	N1	N2	N3	REM	
Manual	Awake	222	0	0	0	45	0.83
	N1	63	63	153	9	144	0.14
	N2	24	42	1431	63	318	0.76
	N3	0	0	252	1125	27	0.80
	REM	6	96	54	0	963	0.86
Average							0.68

TABLE IV: Sensitivity values of AQSSA-SE, Wavelet-based method and single channel SSA-SE

	Individual stages					Average
	Awake	N1	N2	N3	REM	
AQSSA-SE	0.83	0.14	0.76	0.80	0.86	0.68
Wavelet	0.80	0.14	0.60	0.66	0.72	0.58
SSA-SE	0.61	0.06	0.63	0.44	0.84	0.52

TABLE V: Epochs agreement and Cohen's kappa coefficient for AQSSA-SE, Wavelet-based method and single channel SSA-SE

	Epoch-by-epoch		Cohen's kappa	
AQSSA-SE	0.74	0.65	(Substantial agreement)	
Wavelet	0.62	0.51	(Moderate agreement)	
SSA-SE	0.59	0.44	(Moderate agreement)	

Statistical description of the training and validation groups is summarised as:

	Epochs	Awake	N1	N2	N3	REM
Training	4800	6.0%	8.1%	37.3%	24.9%	23.7%
Validation	5100	5.2%	8.5%	36.8%	27.5%	22.0%

In order to compare the AQSSA-SE with a related method, the M-SVM was provided with the descriptors obtained from both AQSSA-SE and Morlet wavelet. Additionally, to illustrate the benefits of multichannel analysis, we performed classification using a single channel SSA (SSA-SE) on central channel C4-A1. Table III shows the confusion matrix of five-stage manual scoring versus automated scoring via AQSSA-SE.

The sensitivity of automated scoring corresponding to each stage and the average sensitivity for AQSSA-SE, SSA-SE and wavelet methods are given in Table IV.

In addition, the epoch-by-epoch agreement and Cohen's kappa coefficient were computed based on the total of 5100 epochs from test subjects, see Table V.

As an illustration example, Fig. 6 illustrates both manual and automatic scoring obtained for the first 350 epochs of an individual. Moreover, Fig. 7 shows the 3D spectrogram of the corresponding data for the first 15 minutes. Observe that the spectrum is in agreement with the literature on sleep analysis [1], [31], i.e. the power of the brain rhythms in decreasing order are from delta, then theta, alpha, **sigma** (spindles) and to beta respectively. For example, the power of delta is greater than that of theta. Furthermore, this figure illustrates the temporal information of different bands, such that alpha and beta are dominant at the start of the data. Over time, the spectrum becomes more influenced by lower frequency rhythms, i.e. delta and theta.

D. Discussion

In contrast to our previous work which detected two stages of sleep using only delta and alpha, in this **work**, our proposed AQSSA-SE can track five subbands delta, theta,

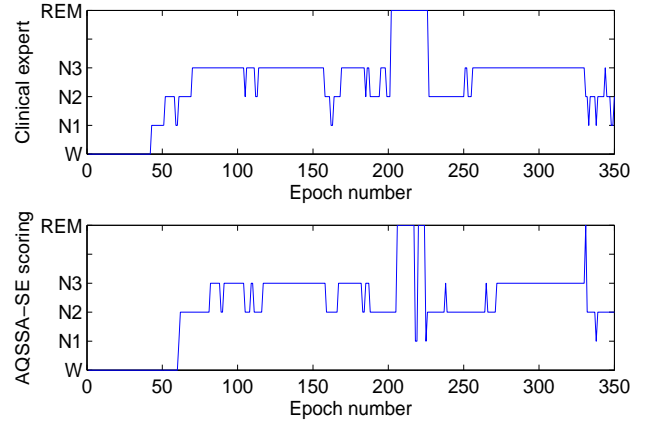


Fig. 6: Representative five-stage sleep classification of the first 350 epochs obtained using manual scoring (top) and AQSSA-SE (bottom).

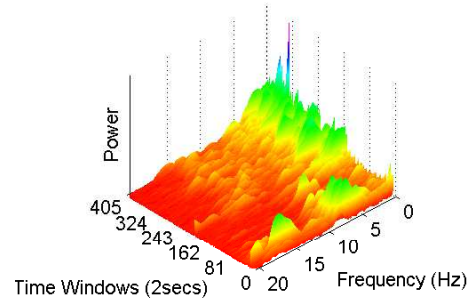


Fig. 7: 3D Spectrogram (Time-frequency-power)

alpha, **sigma**, and beta. Our results demonstrated that the fully automated EEG analysis using our proposed method provided a five-stage scoring system exhibiting 74% agreement with a manually scored standard reference. The AQSSA-SE was more than 0.75 sensitive for detection of wakefulness, N2, N3 and REM stages. Moreover, the kappa coefficient was 65% which indicated substantial level of agreement between manual and automated scorers; while, using the same statistical descriptors, the wavelet-based classification and single channel SSA provided moderate classifications with $\kappa = 51\%$ and $\kappa = 44\%$ respectively. Note that the proposed AQSSA-SE involved an inherent filtering process and this may explain the **higher sensitivities** of its results. Another advantage of AQSSA-SE over other methods such as wavelet and SSA-SE is that it is multi-dimensional and it can exploit the spatial information and the correlation between the channels to improve the results. Moreover, it is possible to extract higher number of sources from a limited number of electrodes which addresses the under-determined case scenarios.

Most of the automatic sleep scoring methods rely on all PSG information including EEG, EOG and EMG channels. However, in this work we focused on four EEG channels. Since REM identification includes the presence of rapid eye movements and muscle atonia [20] and N1 is a transition stage between wakefulness and asleep, differentiation of REM from wakefulness and N1 is difficult without using any EMG or EOG channel. To overcome this limitation, we have utilised the ratio $\frac{\text{beta} \times \text{theta}}{\text{delta}}$ as a specific marker for discrimination of REM stage [25]. Thus, the automated scoring provided sensitivities 0.83 and 0.86 for wakefulness and REM respectively, see **Table IV**. However, the discrimination between N1 and REM

was still challenging due to their similarity in EEG pattern [24] which caused low sensitivity of stage N1, see Table IV.

Remark#10: The aim of this paper is to propose a novel multivariate EEG analytic framework, which can be useful in rehabilitation settings. As an application, we have considered sleep analysis and based our on the work BL nights of healthy subjects only. Extension of this work to a more diverse set of subjects (such as those suffering from apnea) is beyond the scope of our work.

VIII. CONCLUSION

A quaternion extension of SSA has been introduced for the first time to deal with general three- and four-dimensional EEG processing. It has been shown that quaternion SSA has generalised our previous work in single channel processing to multi-channel processing in the context of under-determined source separation. Our previous work [8], however, required an optimisation step (17), which limited the scope on the automatic grouping. Instead, our proposed method AQSSA-SE determines the grouping of the selected elementary matrices in (3) by measuring the correlation between the reconstructed signal and the desired frequency bands. That allowed for all the sleep stages to be detected in an automated way. The robustness of our approach has been demonstrated by the low SNR (down to -10 dB) in synthetic EEG datasets. In the context of sleep EEG, our proposed method was used to determine the statistical descriptors for five-stage sleep classification. The proposed method provided 74% epoch agreement between the automated and manual scoring carried out by the Sleep Centre of University of Surrey. In addition, the proposed method has a great potential to be used in various multivariate biomedical applications, such as BCI or gait analysis for rehabilitation purposes.

ACKNOWLEDGMENT

The authors would like to thank Prof. D.-J. Dijk from Surrey Sleep Research Centre for his technical support and providing the relevant dataset.

REFERENCES

- [1] S. Sanei, *Adaptive Processing of Brain Signals*. John Wiley & Sons, 2013.
- [2] H. Hassani, R. Mahmoudvand, and M. Zokaei, "Separability and window length in singular spectrum analysis," *Comptes rendus mathematique*, vol. 349, no. 17, pp. 987–990, 2011.
- [3] N. Golyandina, V. Nekrutkin, and A. A. Zhigljavsky, *Analysis of time series structure: SSA and related techniques*. CRC Press, 2010.
- [4] S. Sanei, T. Lee, and V. Abolghasemi, "A new adaptive line enhancer based on singular spectrum analysis," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 2, pp. 428–434, 2012.
- [5] F. Ghaderi, H. Mohseni, and S. Sanei, "Localizing heart sounds in respiratory signals using singular spectrum analysis," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 12, pp. 3360–3367, 2011.
- [6] L. Spyrou, Y. Blokland, J. Farquhar, and J. Bruhn, "Singular spectrum analysis as a preprocessing filtering step for FNIRS brain computer interfaces," *Eur. Signal Process. Conf. (EUSIPCO)*, pp. 46–50, 2014.
- [7] T. K. Lee, S. S. Gan, S. Sanei, and S. Kouchaki, "Assessing rehabilitative reach and grasp movements with singular spectrum analysis," *Eur. Signal Process. Conf. (EUSIPCO)*, pp. 1–5, 2013.
- [8] S. Kouchaki, S. Sanei, E. Arbon, and D. Dijk, "Tensor based singular spectrum analysis for automatic scoring of sleep EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 1, pp. 1–9, Jan 2015.
- [9] S. Javidi, C. C. Took, and D. Mandic, "Fast independent component analysis algorithm for quaternion valued signals," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 1967–1978, 2011.
- [10] S. Enshaeifar, S. Sanei, and C. C. Took, "Singular spectrum analysis for tracking of P300," *IEEE Int. Conf. Neural Netw. (IJCNN)*, pp. 502–506, 2014.
- [11] S. Enshaeifar, S. Sanei, and C. C. T ook, "An eigen based approach for complex valued forecasting," *IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, pp. 6014–6018, 2014.
- [12] C. Park, C. C. Took, and D. Mandic, "Augmented complex common spatial patterns for classification of noncircular EEG from motor imagery tasks," *IEEE Trans. Neural Syst. and Rehabil. Eng.*, vol. 22, no. 1, pp. 1–10, 2014.
- [13] E. Niedermeyer, "Alpha rhythms as physiological and abnormal phenomena," *Psychophysiology*, vol. 26, no. 1-3, pp. 31–49, 1997.
- [14] R. Mistlberger, B. Bergmann, and A. Rechtschaffen, "Relationships among wake episode lengths, contiguous sleep episode lengths, and electroencephalographic delta waves in rats with suprachiasmatic nuclei lesions," *Sleep*, vol. 10, no. 1, pp. 12–24, 1987.
- [15] C. C. Took and D. P. Mandic, "Augmented second-order statistics of quaternion random signals," *Signal Process.*, vol. 91, no. 2, pp. 214 – 224, 2011.
- [16] N. Le Bihan and J. Mars, "Singular value decomposition of quaternion matrices: A new tool for vector-sensor signal processing," *Signal Process.*, vol. 84, no. 7, pp. 1177–1199, 2004.
- [17] C. C. Took, G. Strbac, K. Aihara, and D. Mandic, "Quaternion-valued short-term joint forecasting of three-dimensional wind and atmospheric parameters," *Renew. Energy*, vol. 36, no. 6, pp. 1754–1760, 2011.
- [18] J. Navarro-Moreno, R. M. Fernandez-Alcala, and J. C. Ruiz-Molina, "A quaternion widely linear series expansion and its applications," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 868–871, 2012.
- [19] L. De Gennaro and M. Ferrara, "Sleep spindles: an overview," *Sleep Med. Rev.*, vol. 7, no. 5, pp. 423–440, 2003.
- [20] American Academy of Sleep Medicine, C. Iber *et al.*, "The aasm manual for the scoring of sleep and associated events: rules, terminology and technical specifications," *Amer. Academy of Sleep Med.*, 2007.
- [21] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *The Roy. Soc. A: Math., Phys. and Eng. Sci.*, vol. 466, pp. 1291–1302, 2010.
- [22] Z. J. Koles and A. C. Soong, "EEG source localization: implementing the spatio-temporal decomposition approach," *Electroen. Clin. Neuro.*, vol. 107, no. 5, pp. 343–352, 1998.
- [23] M. V. Yeo, X. Li, K. Shen, and E. P. Wilder-Smith, "Can svm be used for automatic eeg detection of drowsiness during car driving?" *Safety Sci.*, vol. 47, no. 1, pp. 115–124, 2009.
- [24] T. Penzel and R. Conradt, "Computer based sleep recording and analysis," *Sleep Med. Rev.*, vol. 4, no. 2, pp. 131–148, 2000.
- [25] C. Berthomier, X. Drouot, M. Herman-Stoïca, P. Berthomier, J. Prado, D. Bokar-Thire, O. Benoit, J. Mattout, and M.-P. d'Ortho, "Automatic analysis of single-channel sleep eeg: validation in healthy individuals," *Sleep*, vol. 30, no. 11, p. 1587, 2007.
- [26] A. H. Khandoker, M. Palaniswami, and C. K. Karmakar, "Support vector machines for automated recognition of obstructive sleep apnea syndrome from eeg recordings," *IEEE Trans. Inf. Technol. Biomedicine*, vol. 13, no. 1, pp. 37–48, 2009.
- [27] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 2, pp. 205–211, 2007.
- [28] E. Sitnikova, A. E. Hramov, A. A. Koronovsky, and G. van Luijtelaar, "Sleep spindles and spike-wave discharges in EEG: their generic features, similarities and distinctions disclosed with fourier transform and continuous wavelet analysis," *Neurosci. Methods*, vol. 180, no. 2, pp. 304–316, 2009.
- [29] C. Bolduc, A.-M. Daoust, lyse Limoges, C. M. Braun, and R. Godbout, "Hemispheric lateralization of the EEG during wakefulness and REM sleep in young healthy adults," *Brain and Cognition*, vol. 53, no. 2, pp. 193 – 196, 2003.
- [30] J. Virkkala, J. Hasan, A. Väri, S.-L. Himanen, and K. Müller, "Automatic sleep stage classification using two-channel electro-oculography," *J. Neurosci. methods*, vol. 166, no. 1, pp. 109–115, 2007.
- [31] T. Andrillon, Y. Nir, R. J. Staba *et al.*, "Sleep spindles in humans: insights from intracranial EEG and unit recordings," *Neurosci.*, vol. 31, no. 49, pp. 17 821–17 834, 2011.