Journal of Neuroscience Methods, 2017. [Preprint] June 3, 2017

Multichannel Sleep Spindle Detection using Sparse Low-Rank Optimization

Ankit Parekh[a,*], Ivan W. Selesnick[b], Ricardo S. Osorio[c], Andrew W. Varga[d], David M. Rapoport[d], Indu Ayappa[d]

[a] *Dept. of Electrical and Computer Engineering, College of Engineering, University of Iowa*
[b] *Dept. of Electrical and Computer Engineering, Tandon School of Engineering, New York University*
[c] *Center for Brain Health, Department of Psychiatry, School of Medicine, New York University*
[d] *Division of Pulmonary, Critical Care and Sleep Medicine, Icahn School of Medicine at Mount Sinai*

Abstract

Background: Automated single-channel spindle detectors, for human sleep EEG, are blind to the presence of spindles in other recorded channels unlike visual annotation by a human expert.
New Method: We propose a multichannel spindle detection method that aims to detect global and local spindle activity in human sleep EEG. Using a non-linear signal model, which assumes the input EEG to be the sum of a transient and an oscillatory component, we propose a multichannel transient separation algorithm. Consecutive overlapping blocks of the multichannel oscillatory component are assumed to be low-rank whereas the transient component is assumed to be piecewise constant with a zero baseline. The estimated oscillatory component is used in conjunction with a bandpass filter and the Teager operator for detecting sleep spindles.
Results and comparison with other methods: The proposed method is applied to two publicly available databases and compared with 7 existing single-channel automated detectors. F_1 scores for the proposed spindle detection method averaged 0.66 (0.02) and 0.62 (0.06) for the two databases, respectively. For an overnight 6 channel EEG signal, the proposed algorithm takes about 4 minutes to detect sleep spindles simultaneously across all channels with a single setting of corresponding algorithmic parameters.
Conclusions: The proposed method attempts to mimic and utilize, for better spindle detection, a particular human expert behavior where the decision to mark a spindle event may be subconsciously influenced by the presence of a spindle in EEG channels other than the central channel visible on a digital screen.

Keywords: Sleep EEG, spindle detection, multichannel signal processing, sparse signal, convex optimization

1. Introduction

Sleep spindles are short rhythmic oscillations visible on an electroencephalograph (EEG) during non-rapid eye movement (NREM) sleep. The center frequency of sleep spindles is between 11 and 16 Hz [59]. The duration of sleep spindles is defined to be at least 0.5 seconds, with some studies imposing an upper limit on their duration to 3 seconds [67]. Sleep spindles reflect a heritable set of traits which is implicated in both sleep regulation and normal cognitive functioning [39]. Recent studies have linked spindle density (number of spindles per minute), duration, amplitude and peak frequency of spindles to memory consolidation during sleep [32, 15], cognition in schizophrenia patients [39, 66], brain dysfunction in obstructive sleep apnea [14] and biomarkers for Alzheimer's disease [69]. As a result, understanding the characteristics of sleep spindles is a key step in studying their relation to several neuropsychiatric diseases.

Traditionally, sleep spindles are annotated in clinics using visual heuristics: number of peaks or bumps of the EEG signal are counted within a specified time window. Not only is this process subjective and time-consuming, but it is also prone to errors. Moreover, visual inspection underscores the fine details of putated spindles [52]. In order to reduce the subjectivity of visual detection, it is not uncommon for studies to utilize more than one expert for detecting spindles. However, in several cases this leads to a high variability in inter-scorer agreement. The Cohen's κ coefficient for inter-rater agreement in manual scoring usually ranges between 0.46 and 0.89 [61, 43]. As such, the presence of reliable automated spindle detectors may not only reduce the scoring variability [71, 72] but may also aid in complex longitudinal studies that involve studying global or local sleep spindle dynamics [52, 22, 44].

Broadly categorized, there exist two-types of automated sleep spindle detectors for single channel EEG: filtering based and non-linear signal decomposition based. Filtering based approaches vary from basic methods, which utilize a bandpass filter with constant or adaptive thresholds, to advanced methods that use time-frequency information along with bandpass filtering. Most of the filtering based methods involve pre-processing of the desired channel of the EEG (usually a central channel) for arti-

*Corresponding author. Email address: ankit.parekh@nyu.edu
Source Code available at `https://github.com/aparek/mcsleep.git`
`https://dx.doi.org/10.1016/j.jneumeth.2017.06.004`

fact removal [36]. One of the first automated detectors to be proposed used a bandpass filter in conjunction with an amplitude threshold [57]. This idea is still the basis of a majority of the bandpass filtering-based automated detectors [68, 23, 40, 31, 16, 33]. Advanced methods utilizing time-frequency information either use a wavelet transform [37, 2, 4, 30, 64, 3] or a short-time Fourier transform (STFT) [21, 47, 25] with adaptive thresholding to detect spindles. Several machine-learning based spindle detectors and sleep staging algorithms have also been proposed for single channel EEG [1, 35].

Non-linear signal decomposition based methods [50, 51, 37, 27] attempt to separate the non-rhythmic transients or artifacts from sinusoidal spindle-like oscillations in the single channel sleep EEG. These methods make use of the differing morphological aspects [60] of the transients and spindles to overcome the drawbacks of filtering and Fast Fourier Transform (FFT) based techniques [53]. As an another example, Gilles et. al considered the removal of ballistocardiogram (BCG) artifacts from EEG using low-rank and sparse decomposition [34]. In addition to these morphological component analysis (MCA) based methods, independent component analysis (ICA) and principal component analysis (PCA) have also been used to detect spindles for single channel EEG [5]. However, note that ICA assumes linearity and stability of the mixing process along with statistical independence of input sources [28].

1.1. Motivation

Automated spindle detectors that consider only a single channel are blind to the presence of spindles in other recorded channels. Such a spindle detection mechanism may not be in concordance with the way spindles are annotated visually. The American Academy of Sleep Medicine (AASM) manual recommends using F4, C4 and O2 channels (or alternatively Fz, Cz and C4) of the recorded EEG with F3, C3 and O1 as backup channels [59] for scoring of sleep and associated events. As such, while annotating sleep events, such as spindles, rarely does an expert view a single channel of the EEG in isolation to the other channels. This is certainly the case for studies either looking to characterize individual global sleep spindle density [9] or tracking the propagation of spindles overnight [52, 20]. As a result, it may be possible that the presence of spindles in channels other than the channel of interest subliminally influences the experts' decision of marking an event as a spindle.

As an example, consider the 3-channel EEG shown in Fig. 1. The experts visually annotated the presence of a spindle at approx. 26 seconds. While it is suggested that only the central channel was used for annotating spindles [23], it can be seen that the spindle at approx. 26 seconds is also present in the frontal and the occipital channels, though with different amplitudes. As such it is highly likely that the decision by a human to mark the presence of a spindle at approx. 26 seconds in the central channel is reinforced by its presence in other channels if they are

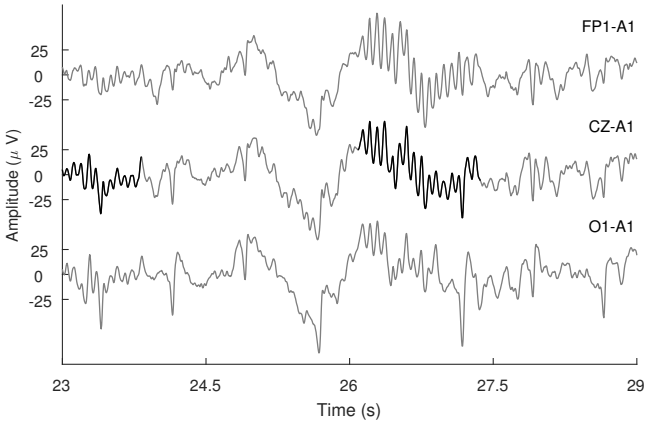

Figure 1: An example of a 3 channel scalp EEG from DREAMS [24] database. Experts annotated two spindles (shown in bold) in the 6 second excerpt using the central channel. The annotated spindle at 26 seconds has different amplitude in different channels.

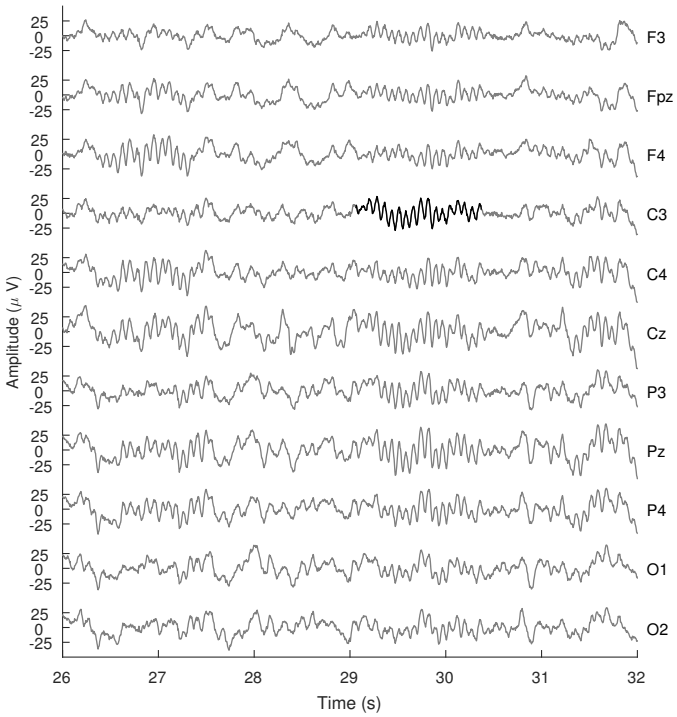

Figure 2: A multichannel channel EEG excerpt from the MASS database [46] (SS2, PSG1, epoch 701). Two experts have visually annotated a spindle at approx. 29 seconds (shown in bold). The spindle at about 26.5 seconds appears predominantly in the right channels.

viewed together on a digital screen. Similar behavior can be seen in the case of the EEG excerpt in Fig. 2 where the experts annotated a spindle at approx. 29.5 seconds. While the degree to which the decision of marking a spindle was influenced by its presence in other channels (if it occurred simultaneously in more than one channel) is an open question out of the scope of this paper, utilizing it can certainly aid in a better design of the automated spindle detectors.

Another motivation for considering multichannel EEG

for studying spindle activity comes from the fact that while single channel detectors may be used to study global spindle activity, their usage comes at a price. Since the amplitude of spindles vary in each channel (see for example Fig. 1), amplitude-based thresholds used by automated detectors need to be tuned separately for each channel, adding to the overall computational complexity. Additionally, the CPU time is multiplied by the number of channels recorded. While this additional computing time may not be significant for the case of basic filtering-with-thresholding methods, it is certainly significant for advanced methods that utilize either time-frequency information or non-linear signal decomposition.

Classifying spindles as either global (occurring across all channels) or local (occurring across a single or group of channels) [15, 12] is difficult using single channel based methods. Spindles that appear on the right channels (F4, C4, and O2 channels in Fig. 2) may be entirely missed by detectors using the left channels (F3, C3 and O1) or vice-versa, which is the case with most detectors [67]. In fact, most detectors utilize at the most two channels [45, 53] possibly from the same hemisphere (left or right). Thus, for a single channel method to effectively detect global spindle activity it has to be run on channels belonging to the left, right hemispheres and also the midline channels. As an example, consider the spindle at approx. 26.5 seconds in the EEG shown in Fig. 2. This spindle appears prominent in the right channel, while appearing in other channels (such as Pz and Cz) with significantly lower amplitude. As a result, without careful parameter tuning an automated detector may not be able to properly detect the spindles.

1.2. Contribution

The contributions of this paper are two-fold. First, we propose a multichannel transient separation method which decomposes the input multichannel EEG as the sum of a transient and an oscillatory component. Second, we utilize the estimated oscillatory component to detect spindles by using an envelope of the bandpass filtered oscillatory component. Combined, the two contributions of this paper aim to detect spindles across all channels of scalp EEG in a single run with a single parameter-tuning (i.e., the parameter-tuning is independent of the number of recorded EEG channels). To the best of the authors knowledge, an automated method for detecting spindles globally in a single run has not been studied before.

The transient component, in the proposed non-linear signal model, is modeled as a sparse piecewise constant signal, whereas the oscillatory component is assumed to exhibit block-similarity, i.e., the blocks of the multichannel oscillatory component are low-rank arrays. We estimate the two components of the proposed non-linear signal model by formulating an optimization problem posed as the minimization of a convex objective function. We derive a fast matrix-inverse-free algorithm to obtain the solution of the proposed optimization problem. Using an envelope of the bandpass filtered oscillatory component we detect sleep spindles globally (across all channels) or locally (across a single or a group of channels).

The bandpass filter used for the detection of spindles is excited by the presence of transients such as BCG, electrocardiogram (ECG) and other non-rhythmic waveforms. By separating the transients from the input multichannel EEG, the proposed method avoids the unnecessary excitation of the bandpass filter. Thus, the spindle activity appears more prominently than if the bandpass filter was directly applied to the input EEG. Moreover, due to the effective attenuation of non-oscillatory transients, a low-order bandpass filter suffices for sleep spindle detection [51, 50].

1.3. Relation to prior work

Several studies have pointed to the benefit of separating the transients and oscillations sleep spindle detection [51, 50, 20], though few have advocated the use of multichannel EEG [6]. Non-linear signal models for the input EEG have been used to detect sleep spindles directly for single channel [51, 50] and indirectly for multichannel EEG [6]. An ICA based approach was studied for the detection of spindles from multichannel EEG in [55]. For sleep-staging and classification a machine learning approach utilizing multichannel EEG was proposed in [58]. A matching pursuit (MP) based decomposition method was proposed for multichannel EEG to relax the assumptions of ICA [28]. The MP based method attempts to represent the multichannel EEG as a linear combination of atoms or coefficients with respect to a chosen basis. Estimating the atoms, by solving an inverse problem, can enable detection of sleep spindles with user chosen parameters [28]. Similar to the ICA approach, a multichannel Matching Pursuit based method was also proposed for the decomposition of the multichannel EEG signal [62]. An MP-based algorithm using singular value decomposition (SVD) was shown to be able to efficiently learn the different oscillatory waveforms in multichannel EEG [11].

1.4. Organization

The rest of the paper is organized as follows. In Section 2 we detail the notation used throughout the paper and introduce the block low-rank operator. In Section 3, we propose a non-linear signal model for the EEG and show how to estimate its components. Further, we formulate and derive the transient separation algorithm and show how to use the estimated components for detecting sleep spindles globally. In Section 4, we illustrate the proposed transient separation and spindle detection methods and show how to set the parameters for the same. In Section 5, we illustrate the proposed method on publicly available annotated single channel spindle databases and finally conclude in Section 6.

2. Preliminaries

2.1. Notation

We denote vectors and matrices by lower and upper case letters respectively. An n-point signal y is represented by the vector

$$y := [y(0), \ldots, y(n-1)], \quad y \in \mathbb{R}^n. \tag{1}$$

We represent the multichannel signal $X \in \mathbb{R}^{k \times n}$, with k channels as

$$X := [x_1, \ldots, x_k]^T, \quad x_i \in \mathbb{R}^n, \quad i = 1, \ldots, k \tag{2}$$

where $[\cdot]^T$ represents the transpose. The ℓ_1 and ℓ_2 norm of the vector y are defined as

$$\|y\|_1 := \sum_{i=1}^{n} |y(i)|, \qquad \|y\|_2^2 := \sum_{i=1}^{n} |y(i)|^2, \tag{3}$$

and the ℓ_1 norm of the matrix $X \in \mathbb{R}^{k \times n}$ is defined as

$$\|X\|_1 := \sum_{i=1}^{k} \sum_{j=1}^{n} |X_{ij}|. \tag{4}$$

The nuclear norm of the matrix $X \in \mathbb{R}^{m \times n}$ is defined as

$$\|X\|_* := \mathrm{tr}(X^T X) \tag{5}$$

$$= \sum_{i=1}^{m} \sigma_i(X), \tag{6}$$

where $\mathrm{tr}(\cdot)$ represents the trace and $\sigma_i(X)$ is the i^{th} singular value of X.

We define the matrix $D \in \mathbb{R}^{(n-1) \times n}$ as

$$D := \begin{bmatrix} -1 & 1 & & & \\ & -1 & -1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix}. \tag{7}$$

Using the matrix D, the first-order difference of a discrete signal $y \in \mathbb{R}^n$ is given by Dy. The soft-threshold function [26] for $\lambda > 0$ is defined as

$$\mathrm{soft}(x; \lambda) := \begin{cases} x - \lambda \dfrac{x}{|x|}, & |x| > \lambda \\ 0, & |x| \leqslant \lambda, \end{cases} \qquad x \in \mathbb{R}. \tag{8}$$

Note that the soft-threshold function in (8) is applied element wise to a vector with threshold $\lambda > 0$.

The Teager-Kaiser energy operator for a discrete-time signal y denoted by $T(\cdot)$ is defined as

$$[T(y)]_n := y^2(n) - y(n-1) \cdot y(n+1). \tag{9}$$

Note that, unless stated otherwise, applying the Teager operator to a multichannel signal X implies that the Teager operator is applied to the channel mean of X. The Teager operator has been commonly used to obtain an envelop of the bandpass filtered signal for spindle detection [47, 3, 30, 21, 67].

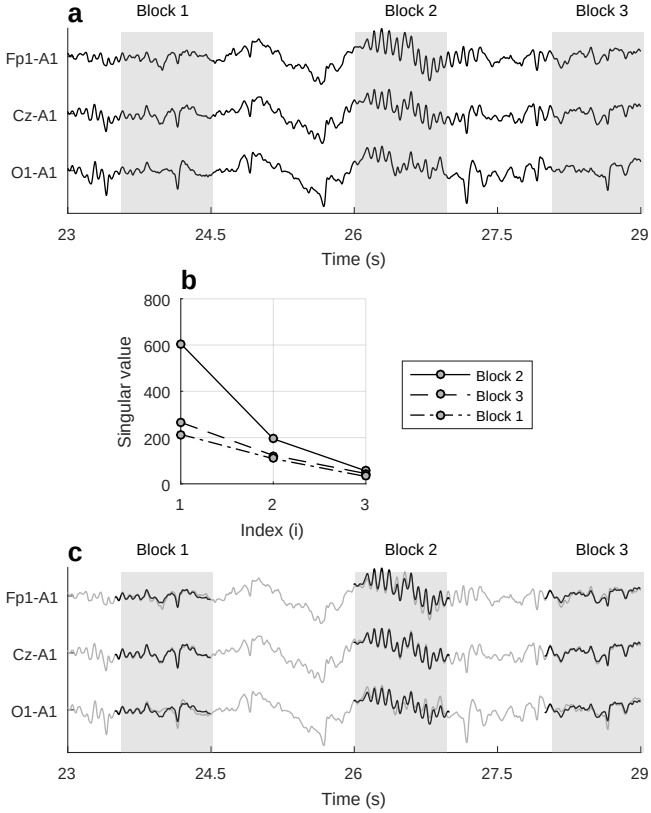

Figure 3: (a) Sample multichannel EEG with multichannel blocks that contain transients and spindles are highlighted (length of each block is 1 second). (b) Singular values of each of the highlighted blocks in (a) are shown. (c) Sample multichannel EEG with the highlighted blocks replaced by their rank-one approximation. Shown in the background is the multichannel EEG from (a).

2.2. Block Low-Rank Operator

In order to extract the rhythmic oscillations (i.e., sleep spindles) from the multichannel EEG, we propose the following sparse optimization framework. Consider a segment of a sample multichannel EEG, as shown in Fig. 3(a) with three blocks highlighted (each of 1 second in length). Expert annotated spindle at approx. 26 seconds overlaps the second multichannel block i.e., Block 2. Figure 3(b) shows the singular values of each of the three blocks. It can be seen in Fig. 3(b) that the Block 2 has larger (in magnitude) singular values than the singular values of Blocks 1 and 3. Moreover, from Fig. 3 it can be seen that the blocks have fact decaying singular values. As such, if one were to approximate each of the three blocks with their low-rank approximations[1], the low-rank approximation for Block

[1]Given an input matrix, a low-rank approximation aims to find a matrix similar to the input matrix but with a reduced rank.

2 would still contain the spindle-like oscillations. Figure 3(c) shows the multichannel EEG from Fig. 3(a) with the blocks replaced by their corresponding rank-one approximation. Furthermore, in the absence of non-spindle like transients the sum of the singular values of all blocks in the multichannel EEG will be approximately equal to the sum of singular values of the spindle-containing blocks. As a result, in order to extract the spindles, it suffices to obtain a low-rank representation for multichannel EEG, once non-rhythmic transients are separated.

We define the operator $\Phi\colon \mathbb{R}^{k\times n} \rightarrow \mathbb{R}^{k\times l\times m}$, which extracts m blocks, each of an even length l, from the $k-$channel input signal as

$$[\Phi(X)]_i := \begin{pmatrix} x_1(i) & \dots & x_1(i+l-1) \\ & \vdots & \\ x_k(i) & \dots & x_k(i+l-1) \end{pmatrix}, \tag{10}$$

for $i = 1, \dots, m$. The operator Φ can be defined with a certain overlap between consecutive blocks. Further, Let the adjoint operator be denoted by $\Phi^T\colon \mathbb{R}^{k\times l\times m} \rightarrow \mathbb{R}^{k\times n}$. The adjoint operator forms the $k-$channel signal by aggregating the m blocks, where by aggregating we mean that the blocks are added in an overlap-add way. Note that in the case of distinct blocks, i.e., no overlap between the blocks, the operator Φ is orthogonal ($\Phi^T\Phi = I$).

In this paper, we use the operator Φ with 50% overlap between blocks of 1 second in length, implemented to obtain perfect reconstruction. As an example, for an EEG signal sampled at 256 Hz, the block length is fixed at 256 samples. In order to perfectly reconstruct the input signal $X \in \mathbb{R}^{k\times n}$ from $\Phi(X)$, we use a diagonal weight matrix $W \in \mathbb{R}^{n\times n}$. Since the blocks are aggregated in an overlap-add way, the samples of signal X that are contained in the overlap occur twice in the signal formed using the adjoint operator Φ^T. As a result, appropriately weighting the samples can lead to perfect reconstruction[2].

An example will help clarify the proposed block low-rank operator. Consider the single channel signal $X = [x_1, x_2, x_3]$. Using a block length of 2 samples with 50% overlap leads to

$$\Phi(X) = \Big\{[x_1, x_2], [x_2, x_3]\Big\}. \tag{11}$$

Reconstructing the signal from the individual blocks by overlapping and adding we get

$$\mathcal{O}\big(\Phi(X)\big) = [x_1, 2x_2, x_3], \tag{12}$$

where $\mathcal{O}(\cdot)$ defines the overlap-add operator. Note that by 'overlap-add' we imply that the individual blocks of size $k \times l$ are overlapped and added to construct a multichannel signal of size $k \times n$. In order to achieve perfect reconstruction, i.e., $\Phi^T(\Phi(X)) = X$, we use the weight matrix $W \in \mathbb{R}^{3\times 3}$, given by

$$W = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{13}$$

and define

$$\Phi^T(\Phi(X)) = \mathcal{O}\big(\Phi(X)\big) \cdot W \tag{14}$$

$$= X. \tag{15}$$

Note that the weight matrix W associated with the operator Φ with 50% overlap can be pre-computed based on the input signal length and the user chosen block length. In particular, for an input signal $X \in \mathbb{R}^{k\times n}$, $n > 3$, the diagonal weight matrix associated with the operator Φ with 50% overlap and an even block length of l is given by

$$\mathsf{diag}(W) = \big\{\underbrace{1, \dots, 1}_{l/2}, \underbrace{1/2, \dots, 1/2}_{n-l}, \underbrace{1, \dots, 1}_{l/2}\big\}. \tag{16}$$

A suitable optimization problem for estimating the oscillatory component with a block low-rank structure is given by

$$C^* := \arg\min_C \left\{ \frac{1}{2}\|Y - \Phi^T(C)\|_2^2 + \lambda \sum_{i=1}^{m} \|c_i\|_* \right\}, \tag{17}$$

where $C = [c_1 \dots, c_m]$, $c_i \in \mathbb{R}^{k\times l}$, $C^* \in \mathbb{R}^{k\times l\times m}$, and $\lambda > 0$ is the regularization parameter. Note that the optimization problem in (17) estimates the blocks c_i from which the multichannel signal S can be calculated using Φ^T (i.e., $S = \Phi^T(C^*)$). The optimization problem in (17) is a sum of convex functions (nuclear norm) and a strictly convex function (ℓ_2 norm squared) and hence is a convex optimization problem. As a result, well developed principles of convex optimization can be leveraged to obtain a global minimum [49].

The solution to the optimization problem in (17) can be obtained using the iterative shrinkage/thresholding algorithm (ISTA) [7] and its variants. The ISTA algorithm, for the optimization problem in (17), entails soft-thresholding the singular values of each block of the multichannel signal Y. For an overnight multichannel EEG signal, roughly 30000 blocks of length 1 second are obtained using the operator Φ and as such the ISTA algorithm involves computing 30000 singular value decompositions (SVD). However, since the number of channels is much less than the length of the block, we need compute only the first k (number of channels) singular values and their corresponding left and right singular vectors. As an example, for the multichannel EEG signal shown in Fig. 3 or for the one in Fig. 2 it suffices to compute only the first 3 or 11 singular values respectively.

[2]Note that using a generic amount of overlap does not guarantee perfect reconstruction i.e., $\Phi^T\Phi \neq I$.

3. Sleep Spindle Detection for Multichannel EEG

3.1. Non-linear signal model

We propose the following non-linear signal model for the multichannel EEG denoted by Y:

$$Y := X + S + W, \qquad Y, X, S, W \in \mathbb{R}^{k\times n}, \tag{18}$$

where X represents the transient component, S represents the oscillatory component and W represents additive white Gaussian noise (AWGN) (i.e., $W \sim \mathcal{N}(0, \sigma)$). We assume that the transient component X is sparse and piece-wise constant and the blocks of the oscillatory component S are low-rank as described in Sec. 2.2.

The signal model presented in this paper contains certain similarities to the one presented in [51], in particular, the transient component is modeled in a similar way. Moreover, in both models, sparsity of the block structure of the spindle component is exploited. Although, the properties of the spindle component presented in this paper are different, the overall theme of non-linear signal models presented in this paper, in [51] and in [50] is similar: the input EEG signal is modeled as a sum of transient and oscillatory components.

3.2. Estimating Transient and Oscillatory Components

In order to detect spindles, we first estimate the transient and the oscillatory components in the proposed signal model (18) from the recorded multichannel EEG. To this end, we utilize a sparse optimization framework and propose to solve the following objective function

$$\{X^*, C^*\} = \arg\min_{X,C}\Bigg\{\frac{1}{2}\|Y - X - H(C)\|_2^2 + \lambda_0 \sum_{i=1}^{k}\|x_i\|_1 + \lambda_1 \sum_{i=1}^{k}\|Dx_i\|_1 + \lambda_2 \sum_{i=1}^{m}\|c_i\|_*\Bigg\}, \tag{19}$$

where $X = [x_1, \ldots, x_k]$, $x_i \in \mathbb{R}^n$, $C = [c_1, \ldots, c_m]$, $c_i \in \mathbb{R}^{k\times l}$ and $\lambda_i > 0$ are the regularization parameters. Let $H = \Phi^T$, where Φ is the block low-rank operator as discussed in Sec. 2.2. Recall that D is the first-order difference matrix, as defined in (7), and C is the coefficient array obtained using the operator Φ as defined in (10).

The proposed objective function seeks the optimal solution X^* which is sparse and piecewise constant. The ℓ_1 norm on X penalizes non-sparse solutions and the ℓ_1 norm on Dx_i, for $i = 1, \ldots, k$, penalizes non piecewise constant solutions. These two penalties combined are generally termed as the 'fused-lasso' penalty [63, 48] and have been shown to model the transient component [51] with relative accuracy.

The nuclear norm on each of the coefficients c_i in (19) penalizes solutions C^* that do not exhibit the block low-rank property as described in Sec. 2.2. Using the solution C^* from the optimization problem in (19), we estimate the oscillatory component using the operator $H = \Phi^T$, i.e.,

$$S = \Phi^T(C^*). \tag{20}$$

The estimate for the oscillatory component S can then used to detect sleep spindles.

Algorithm 1 McSleep algorithm for solution to (19)

inputs: $Y \in \mathbb{R}^{k\times n}, \mu > 0, \lambda_i \geqslant 0, \quad i = 0, 1, 2.$
initialize: $D_1, X, D_2, C \leftarrow 0$
repeat
  $f_1 \leftarrow \frac{1}{\mu}Y + X + D_1$
  $f_2 \leftarrow \frac{1}{\mu}H^T Y + C + D_2$
  $U \leftarrow f_1 - \frac{1}{\mu+2}\big(f_1 + Hf_2\big)$
  $V \leftarrow f_2 - \frac{1}{\mu+2}H^T\big(f_1 + Hf_2\big)$
  $x_i \leftarrow \mathbf{soft}\big(\mathbf{tvd}(u_i - d_{(1,i)}, \lambda_1/\mu), \lambda_0/\mu\big)$
  $[\tilde{U}, \tilde{\Sigma}, \tilde{V}] \leftarrow \mathbf{svd}\big(v_i - d_{(2,i)}\big)$
  $c_i \leftarrow \tilde{U} \cdot \mathbf{soft}(\tilde{\Sigma}, \lambda_2/\mu) \cdot \tilde{V}^T$
  $D_1 \leftarrow D_1 - (U - X)$
  $D_2 \leftarrow D_2 - (V - C)$
until convergence
$X = [x_1, \ldots, x_k], \quad X \in \mathbb{R}^{k\times n}$
$S = \Phi^T(C), \quad C \in \mathbb{R}^{k\times l\times m}, \quad C = [c_1 \ldots, c_m]$
return X, S

3.3. Transient Separation Algorithm

We develop a fast iterative algorithm to obtain the optimal solution for X^* and C^* using the proposed objective function (19). Note that the objective function proposed is convex and hence well developed theory of convex optimization algorithms can be leveraged to obtain the optimal solution. We apply Douglas-Rachford splitting [18] to solve (19), which results in an instance of the alternating direction method of multipliers (ADMM) method. The convergence of the iterative ADMM algorithm is guaranteed for the proposed objective function (19) under suitable assumptions [29, 10].

We write the objective function (19) using variable-splitting as

$$\arg\min_{X,U,C,V}\Bigg\{\frac{1}{2}\|Y - U - H(V)\|_2^2 + \lambda_0 \sum_{i=1}^{k}\|x_i\|_1 + \lambda_1 \sum_{i=1}^{k}\|Dx_i\|_1 + \lambda_2 \sum_{i=1}^{m}\|c_i\|_*\Bigg\} \quad \text{such that} \quad U = X, \quad V = C, \tag{21}$$

where $U \in \mathbb{R}^{k\times n}$ and $V \in \mathbb{R}^{k\times l\times m}$. Using the scaled augmented Lagrangian, minimizing (21) results an iterative procedure consisting of three sub-problems and Lagrange multiplier update equations. The following are the three sub-problems:

$$U, V \leftarrow \arg\min_{U,V}\Bigg\{\frac{1}{2}\|Y - \big(U + H(V)\big)\|_2^2 + \frac{\mu}{2}\|U - X - D_1\|_2^2 + \frac{\mu}{2}\|V - C - D_2\|_2^2\Bigg\}, \quad (22a)$$

$$X^* \leftarrow \arg\min_{X}\Bigg\{\frac{\mu}{2}\|U - X - D_1\|_2^2 + \lambda_0\sum_{i=1}^{k}\|x_i\|_1 + \lambda_1\sum_{i=1}^{k}\|Dx_i\|_1\Bigg\}, \quad (22b)$$

$$C^* \leftarrow \arg\min_{C}\Bigg\{\frac{\mu}{2}\|V - C - D_2\|_2^2 + \lambda_2\sum_{i=1}^{m}\|c_i\|_*\Bigg\}, \quad (22c)$$

where $\mu > 0$ is the Lagrangian step-size parameter and $D_1 \in \mathbb{R}^{k\times n}$ and $D_2 \in \mathbb{R}^{k\times l\times m}$ are Lagrange multipliers.

The first term in sub-problem (22b) can be written as the energy over each channel of U, X and D_1, which leads to

$$X^* \leftarrow \arg\min_{X}\Bigg\{\sum_{i=1}^{k}\frac{\mu}{2}\|u_i - x_i - d_{(1,i)}\|_2^2 + \lambda_0\|x_i\|_1 + \lambda_1\|Dx_i\|_1\Bigg\}, \quad (23)$$

where the terms u_i, x_i and $d_{(1,i)}$, for $i = 1, \ldots, k$ represent the $k-$channels of U, X and D respectively. As a result, we write the sub-problem in (23) over each x_i as

$$x_i^* \leftarrow \arg\min_{x_i}\Bigg\{\frac{\mu}{2}\|u_i - x_i - d_{(1,i)}\|_2^2 + \lambda_0\|x_i\|_1 + \lambda_1\|Dx_i\|_1\Bigg\}, \quad (24)$$

with $X^* = [x_1^*, \ldots, x_k^*]$ and $x_i^* \in \mathbb{R}^n$. The solution to (23), for each x_i^*, is readily obtained by applying the fused-lasso method [63] to each channel of the underlying signal, i.e.,

$$x_i^* = \text{soft}(\text{tvd}(u_i - d_{(1,i)}, \lambda_1/\mu), \lambda_0/\mu), \quad (25)$$

where u_i and $d_{(1,i)}$ are the i^{th} channel of U and D_1 respectively. Note that tvd$(\cdot)$ represents the solution of total variation denoising method [56] obtained using a fast solver [19] and soft$(\cdot)$ represents the soft-thresholding function (8).

As in the case of the sub-problem in (22b), we write sub-problem (22c) as

$$C^* \leftarrow \arg\min_{C}\Bigg\{\sum_{i=1}^{m}\frac{\mu}{2}\|v_i - c_i - d_{(2,i)}\|_2^2 + \lambda_2\|c_i\|_*\Bigg\}, \quad (26)$$

where v_i, c_i and $d_{(2,i)}$ are the i^{th} channel of V, C and D_2 respectively. Alternatively, the sub-problem (26) can be written as the optimization problem over each c_i as

$$c_i^* \leftarrow \arg\min_{c_i}\Bigg\{\frac{\mu}{2}\|v_i - c_i - d_{(2,i)}\|_2^2 + \lambda_2\|c_i\|_*\Bigg\}, \quad (27)$$

with $C^* = [c_1^*, \ldots, c_k^*]$ and $c_i \in \mathbb{R}^{k\times l}$. The solution to (27) is obtained using the singular value thresholding (SVT) algorithm [13], i.e.,

$$[\tilde{U}, \tilde{\Sigma}, \tilde{V}] = \text{svd}(v_i - d_{(2,i)}), \quad (28)$$

$$c_i^* = \tilde{U} \cdot \text{soft}(\tilde{\Sigma}, \lambda_2/\mu) \cdot \tilde{V}^T, \quad (29)$$

where svd$(\cdot)$ represents the singular value decomposition. The SVT algorithm computes the singular values of the input matrix and thresholds them using the soft-threshold function [13].

The objective function in the sub-problem (22a) can be solved exactly using a suitable substitution via the least squares method. Note that the objective function in the sub-problem (22a) is similar to [51, Eq. (20a)], and hence a similar derivation can be used for (22a) in this paper. We detail the derivation in Appendix 9.1. The iterative algorithm for (19) is listed in Algorithm 1 and the MATLAB code is made available online[3].

Figure 5 shows the estimated transient and oscillatory components for a three channel EEG (FP1-A1, CZ-A1, O1-A1) from the DREAMS Database[4]. It can be seen in Fig. 5 that the spindles in the three EEG channels are captured by their respective oscillatory components, whereas the non-oscillatory waveforms are captured by

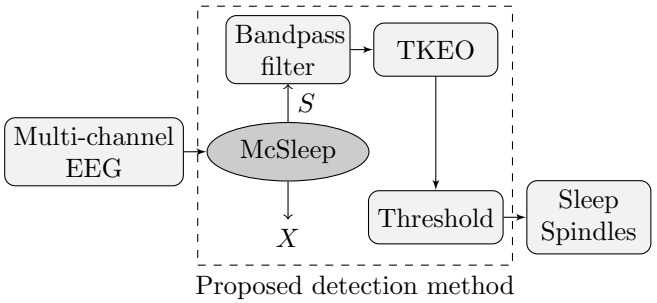

Figure 4: Proposed spindle detection method for multichannel EEG using McSleep as a multichannel transient separation algorithm.

[3] https://github.com/aparek/mcsleep.git

[4] University of MONS - TCTS Laboratory (S. Devuyst, T. Dutoit) and Universite Libre de Bruxelles - CHU de Charleroi Sleep Laboratory (M. Kerkhofs)

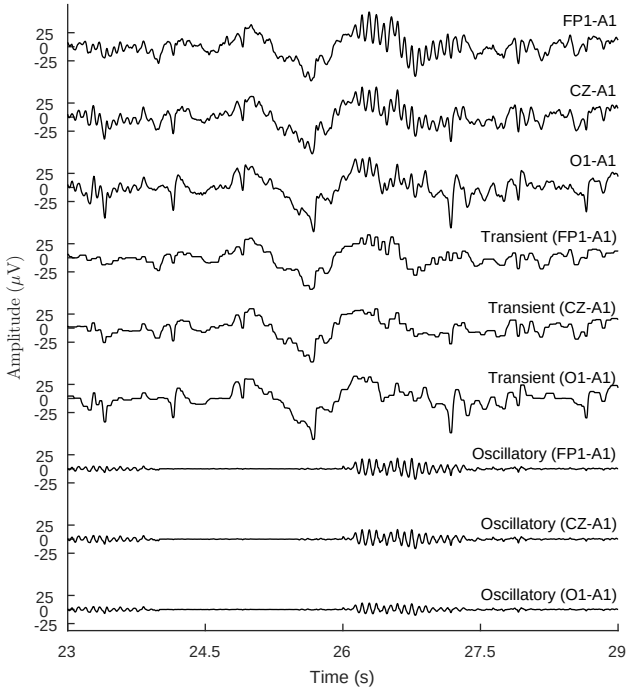

Figure 5: Separation of transients and oscillations using the proposed objective function in (19) for an example EEG segment from DREAMS Database [24]. The transient component is modeled as the sum of a low-frequency signal and a sparse piecewise constant signal.

the transient components. Shown in Fig. 6 is the separation of transients and oscillations for a sample 6-channel EEG from the Montreal Archive of Sleep Studies (MASS) database [46]. The regularization parameters for the preceding examples were found empirically.

3.4. Detection of Spindles Post Separation of Transients

We use the estimated multichannel oscillatory component to detect the sleep spindles. In order to suppress non-spindle like waveforms captured by the oscillatory component, we use a 4^{th} order Butterworth bandpass filter with a user-specified passband. Specifically the bandpass filter is applied to each channel of the estimated oscillatory component. We denote the bandpass filtered oscillatory components as $\text{BPF}(S)$, where S is the oscillatory component.

The usage of the proposed transient-separation algorithm, described in the preceding subsection, allows for the oscillatory activity in the EEG to appear prominently. As a result, post separation of the transients the detection of sleep spindles becomes relatively simpler. We use the Teager Operator, as defined in Sec. 2, to construct an envelope of the oscillatory activity and consequently detect spindles. The Teager operator denoted by $T(\cdot)$, is applied to the channel mean of the multichannel bandpass filtered oscillatory component ($\text{BPF}(S)$) to detect the global spindle activity. Using a constant threshold, we define a binary

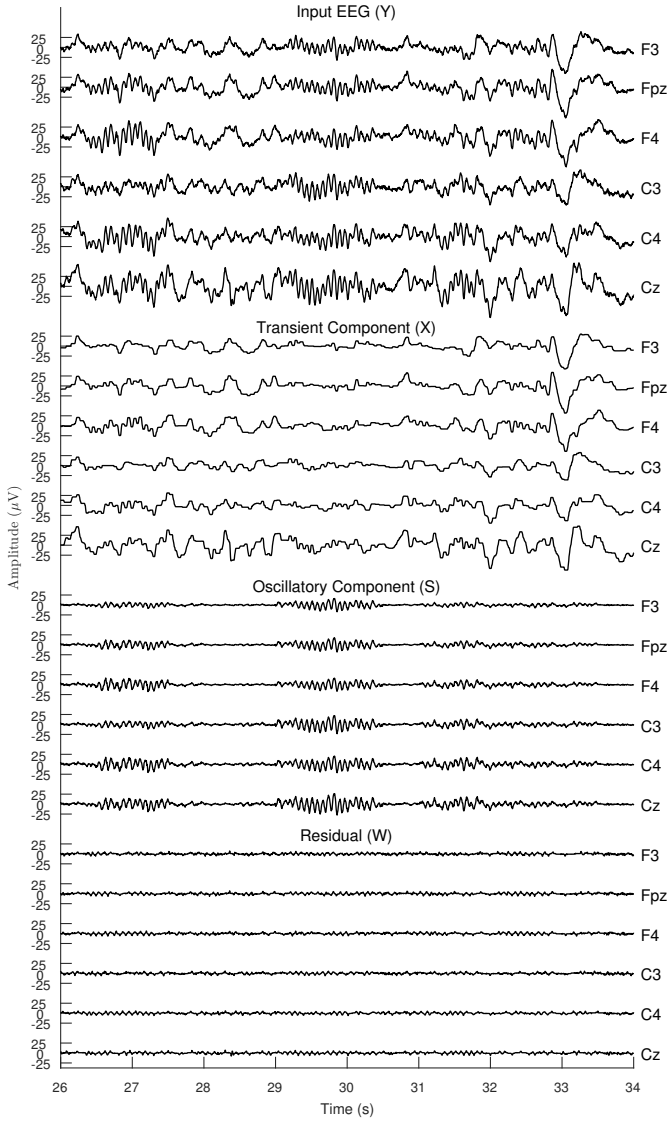

Figure 6: Decomposition of a 6-channel EEG excerpt (Y) into its transient (X) and oscillatory (S) components. Also shown is the the residual W, such that ($Y = X + S + W$). The EEG excerpt is same as the one in Fig. 2.

signal $b_{\text{spindle}}(t)$ as

$$b_{\text{spindle}}(t) := \begin{cases} 1, & T\big(\text{BPF}(S)\big) > c \\ 0, & T\big(\text{BPF}(S)\big) \leqslant c, \end{cases} \tag{30}$$

where 1 denotes a spindle present and 0 otherwise. Figure 4 summarizes the proposed multichannel sleep (McSleep) spindle detection method using the derived transient separation algorithm for sleep EEG.

4. Examples

We illustrate the proposed multichannel sleep spindle detection (McSleep) and compare it to other state-of-the-art automated single-channel based spindle detectors. Recall that the proposed method consists of two parts: sep-

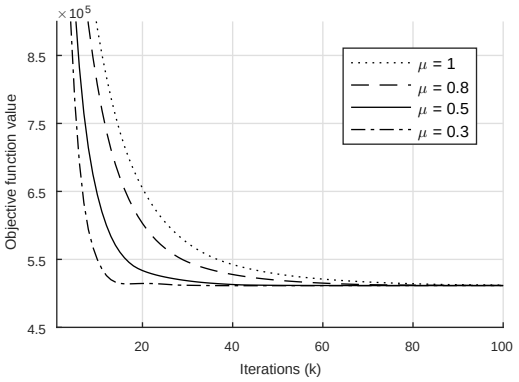

Figure 7: The value of objective function in (19), for each iteration k, is shown for several values of the step-size parameter μ.

aration of transients and oscillations from multichannel EEG and using bandpass filter with Teager operator to detect spindles. Hence, the proposed method can be applied in the following two scenarios. First, when studying global spindle activity using the proposed method the Teager operator is applied to the channel mean of the bandpass filtered oscillatory component. Second, if the user is interested in spindles for only a single channel then the Teager operator, post-estimation of the oscillatory component, can be applied specifically to the desired channel. While the second scenario is a limitation of the proposed method in that only a single channel is considered post estimation of transients, the presence of multiple EEG channels aids in better separation of transients and oscillations for the input EEG. Before we illustrate an example for each of the two scenarios, we provide an overview of the parameters that need to be set for the proposed McSleep spindle detection method.

4.1. Parameters

The proposed spindle detection method requires the user to set several parameters which are either algorithm-specific or task-specific. Algorithm specific parameters are the regularization parameters $\lambda_i \geqslant 0$, for $i = 0, 1, 2.$, in (19) and the step-size μ for the scaled augmented Lagrangian. The regularization parameters influence the sparsity of their respective components. For example, a high value for λ_0, relative to λ_1 and λ_2, enforces the transient component X to be sparse (i.e., with a baseline of zero). Similarly, a high value for λ_2, relative to λ_0 and λ_1 results in the estimated oscillatory component to be of reduced rank. For a sufficiently high λ_2, a rank-one approximation may be obtained.

The values for λ_0, λ_1 and λ_2 were found empirically for the examples that follow. We find that the same λ_0, λ_1 and λ_2 work well for different EEG signals having the same sampling frequency. Thus for a database that contains EEG signals sampled at the same frequency we can preset the values for λ_0, λ_1 and λ_2. In case an EEG contains relatively more transients, the values of λ_0 and λ_1 may be increased proportionally. The step-size parameter μ on the other hand controls the rate of convergence for the proposed algorithm. Note that μ influences the speed at which the algorithm converges and not the solution to which it converges. Figure 7 shows the value of the proposed objective function (19) at each iteration for several values of μ. For the examples that follow and for the experiments in Sec. 5, we fix $\mu = 0.5$. Note that setting the value of μ arbitrarily low or arbitrarily high may affect the convergence of the proposed transient separation algorithm.

The task-specific parameters are the block length for the block low-rank operator H, the overlap between consecutive blocks and the threshold c for the binary vector $b_{\text{spindle}}(t)$. The average duration of a sleep spindle is between 1 and 1.5 seconds [67, 4, 52]. As a result, we set the block length to be fixed at 1 second with a 50% overlap between consecutive blocks. The passband for the Butterworth filter used in this paper is in fact an additional parameter to be set when using the proposed method. We select the passband to be 11 Hz to 16 Hz based on spindle frequency range reported in [23] and the AASM manual [59]. However, increasing number of studies are reporting spindle frequencies to be in a variety of ranges, such as 11–15 Hz [42], 10–17 Hz [67]. As such, there seems to be a lack of consensus among the sleep medicine community regarding the range of spindle frequency.

Recall that the proposed McSleep method detects spindles using a bandpass filter followed by the Teager operator. As a result, studies that may be interested in slow spindles (spindle frequency less than 13 Hz [67]) can set the passband as 10 Hz to 13 Hz or alternatively can set the passband as 13 Hz to 16 Hz for fast spindles. In this manner, the proposed method offers flexibility for the study of spindles, fast and slow alike. Moreover, since the computationally heavy transient separation algorithm needs to be run only once (for a fixed set of regularization parameters), the additional runtime in detecting slow and fast spindles separately is not significant.

Although the list of parameters required to be set by the user for the proposed method is not short, it is worth noting that λ_2 and c are the only parameters that may need to be changed; all other parameters can be fixed for EEG signals that share the same sampling frequency. In Sec. 5.3 we explain how to tune parameters for large databases in a semi-supervised fashion.

4.2. Detecting global and local spindles

We illustrate the proposed McSleep method for detecting global spindle activity in a multichannel EEG. Note that global spindle activity refers to the derivation of a single binary vector denoting the presence or absence of a spindle in a multichannel EEG. We consider a sample segment of the multichannel EEG from the MASS database (Cohort 1, subset 2). For illustration purposes, we consider only three frontal, three central, three parietal and two occipital channels of the multichannel EEG. The EEG

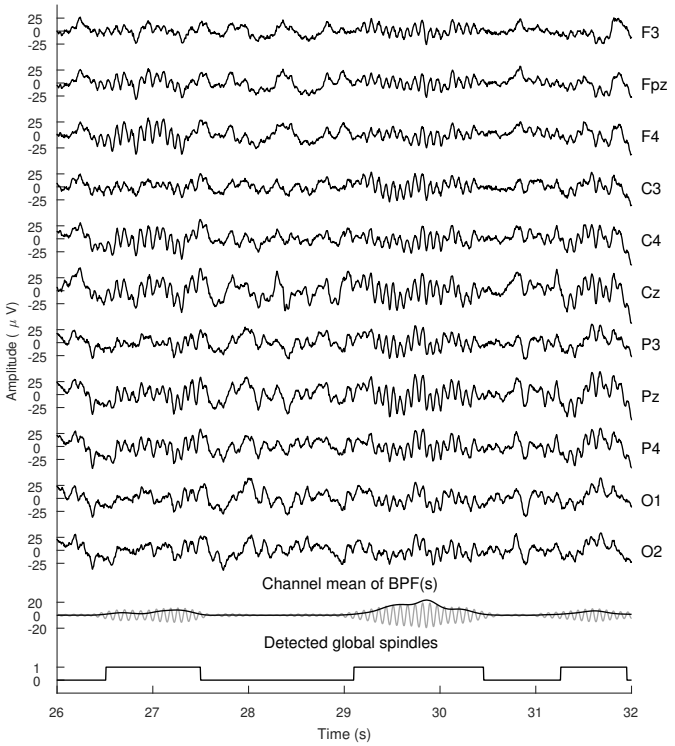

Figure 8: Global Sleep spindle detection on multichannel EEG using proposed McSleep method. Spindles that appear only on the left (alternatively only on the right) channels are also detected by the proposed method.

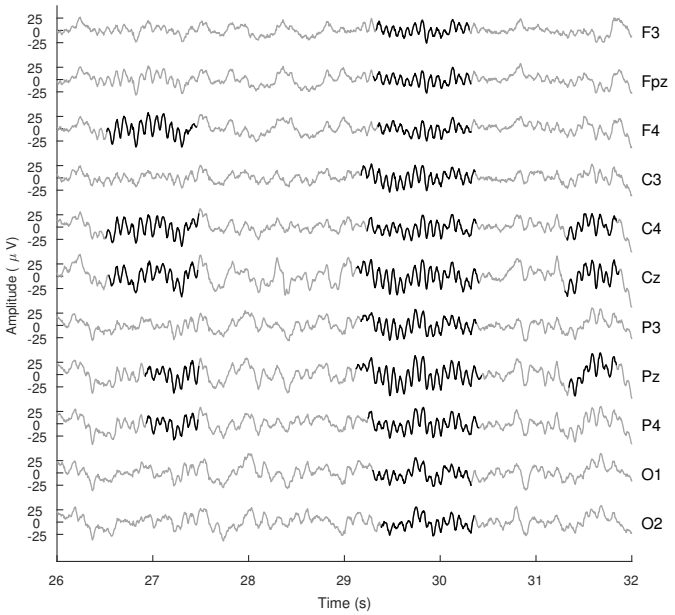

Figure 9: Sleep spindles detected locally (marked in bold) in a single run by applying the proposed McSleep detection method on multiple scalp channels of sleep EEG.

is shown in Fig. 8. Also shown in Fig. 8 is the channel mean of the bandpass-filtered oscillatory component, BPF(s), and its Teager envelope. Note that for visual clarity the channel mean of BPF(s) and its envelope are scaled. The detected spindles are shown using a binary vector at the bottom of Fig. 8.

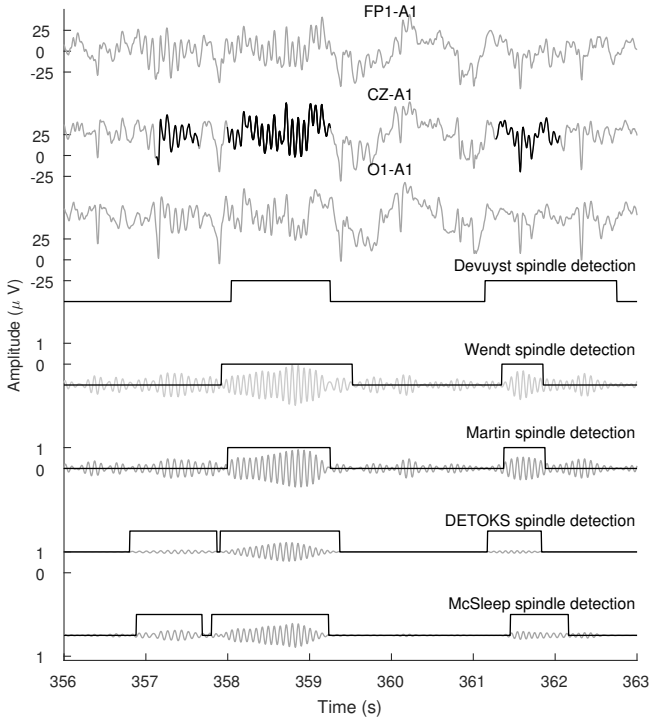

Figure 10: Comparison of proposed McSleep method for single-channel spindle detection with existing spindle detectors. Expert annotated sleep spindles are highlighted in black. Bandpass filter output is also shown in background for several methods.

In order to detect spindles locally (i.e., across a single channel or a group of channels) in a single run, we first run the proposed transient separation algorithm on the input multichannel EEG. We then apply the Teager operator individually to each bandpass filtered oscillatory component using the same threshold value c. Figure 9 shows the spindles detected across the eleven channels from Fig. 8. Note that applying the Teager operator to the channel mean of BPF(S) has the effect of combining the detections from individual channels. As a result, depending on the user, the proposed method can be used to obtain either a single binary vector depicting global spindles or several binary vectors depicting spindles across multiple channels.

4.3. Comparison with existing single-channel spindle detectors

We compare the proposed multichannel sleep (McSleep) spindle detection method, for single channel spindle detection, with the following state-of-the-art automated detectors: Devuyst [23], Wendt [68], Martin [40], and DETOKS [51]. In order to extract spindles for a single channel, we first apply the transient separation algorithm on the multichannel input EEG (FP1-A1, CZ-A1, O1-A1). Then, as noted in the preceding section, post-estimation of the transient and oscillatory component, we apply the Teager operator to the bandpass filtered oscillatory component of the desired channel. Although spindle detection carried out in this manner is an under-utilization of McSleep, the

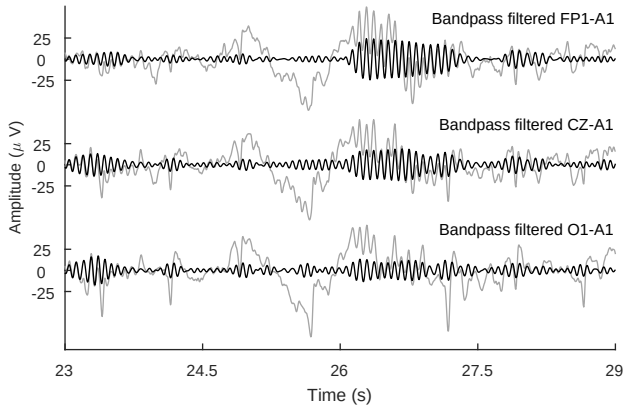

Figure 11: Bandpass filtered sleep EEG using Wendt algorithm [68]. The input EEG is shown in the background. As seen, the presence of transients excites the bandpass filter which may lead to false detections.

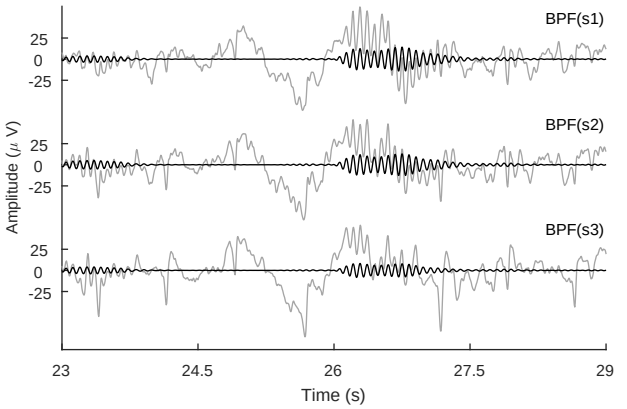

Figure 12: Bandpass filtered oscillatory component estimated using the proposed method. Due to separation of transients, the spindle activity is displayed more prominently.

proposed transient separation method certainly benefits from using a multichannel EEG input. In fact, the more channels present in the input EEG the better the separation between transients and oscillations for the input EEG (see Sec. 2.2).

Figure 10 shows the detection of sleep spindles for an example 3-channel EEG using the Devuyst, Wendt, Martin, DETOKS and the proposed McSleep methods. Also shown in Fig. 10 is the bandpass filter result for the different methods. Due to the absence of the implementation details for the bandpass filter used by Devuyst, we do not show the bandpass filter result in Fig. 10. Note that the experts have annotated three spindles at 357, 358 and 361 seconds visually for the central (CZ-A1) channel only.

The Devuyst, Wendt and Martin detection methods are not able to detect all the three spindles, with the Wendt method detecting a false positive spindle. The DETOKS method does detect all three spindles, but the estimated durations do not closely resemble the expert annotated spindle duration. Note that it is possible to increase the value of the Teager threshold for the DETOKS method to better match the duration of detected spindles. However, it is likely that this will discard previously detected spindles. On the other hand, the proposed McSleep method detects all three spindles and their estimated duration is similar to the expert detection. Moreover, the bandpass filter output for McSleep (only the central channel) shows the spindles much more prominently than the other methods.

Figure 11 shows the bandpass filtered EEG signal using the filter used by the Wendt algorithm [68]. As seen in Fig. 10, and reported in [51], the Wendt method detects false positive spindles due to the presence of transients. In particular, the transients in the sleep EEG excite the bandpass filter and as such the spindle activity does not appear prominent. This leads to the algorithm detecting false positive spindles in areas where non-oscillatory transients are present. It may also generally lead to a high number of false negatives. On the other hand, the proposed method seeks to first separate the transients and then use the estimated oscillatory component to detect spindles, thereby resulting in a much more prominent spindle activity in the bandpass filtered signal as seen in Fig. 12.

5. Illustration of McSleep for single channel spindle detection

To further illustrate the proposed McSleep method we apply it to two online databases and compare the spindle detection results with several state-of-the-art automated single-channel detectors. Specifically we use the DREAMS [23] and the MASS [46] databases which have multichannel EEG recordings and spindles annotated only in the central channel.

5.1. Datasets and existing automated detectors

The DREAMS database[5] was acquired using a 32-channel polygraph (BrainnetTM System of MEDATEC, Brussels, Belgium) [23]. The subjects possessed different pathologies (dysomnia, restless legs syndrome, insomnia and apnoea /hypopnoea syndrome) [24]. The online database provides 8 excerpts of 30 minutes from the whole-night recording. The excerpts contain three EEG channels (FP1-A1, C3-A1 or CZ-A1, and O1-A1), two EOG channels and one submental EMG channel. These excerpts were scored independently by two experts for sleep spindles. Out of the 8 excerpts, we use only 5 of the excerpts as these were scored by both the experts.

The MASS database [46] (Cohort 1, subset 2) contains full night multichannel EEG recordings from 19 healthy young subjects. The overnight recordings contain 19 scalp channels sampled at a frequency of 256 Hz. Of the 19 recorded channels, we use three frontal (F3, F4 and Fz), three central (C3, C4, and Cz) channels for the proposed McSleep method. All the recordings are annotated for

[5] `http://www.tcts.fpms.ac.be/~devuyst/#Databases`

spindles by two experts using the C3 channel and a linked-ear reference. The second expert annotated spindles using broad-band EEG signals (0.35 Hz - 35 Hz) and sigma filtered signals (11 Hz - 17 Hz) similar to [54, 38]. Out of the 19 recordings only 15 were annotated by both the experts and as such we use those 15 recordings. We converted the visual annotations from the EDF+C format to a comma separated value (csv) file using the EDFBrowser software[6]

We use the following state-of-the-art automated detection algorithms: Wendt [68], Martin [40], Bodizs [8], Wamsley [66], Mölle [33], Devuyst [24], and DETOKS[7] [51]. We refer the reader to [67, 50], and [51] for a review and the source code of the detection methods. We base the choice of detection algorithms for comparison on the availability of the source code for each of the algorithms. We apply the existing single-channel detectors to the central channel (C3 or Cz) of both the online databases. For the proposed McSleep method, we run the transient separation method on all the specified channels (3 for DREAMS and 6 for MASS) and apply the Teager operator only to the central (C3 or Cz) bandpass filtered oscillatory channel for detecting spindles.

For both the databases the epochs containing electrode artifacts were identified visually, as in [25, 23], and discarded. The electrode artifacts considered were: lead movements and other body movement artifacts that result in abnormal jump in the amplitude of the EEG. Furthermore, we discard all detected spindles which are either less than 0.5 seconds or greater than 3 seconds [67].

5.2. Measure of performance

We use the expert detection (visually annotated spindles using the central channel) as a gold standard for evaluating the performance of the automated detection algorithms. We use the by-sample method of analysis, as described in [67] with the gold standard as the 'union' of the detections by both the experts. For the 'by-sample' rule, a time sample of the EEG is marked as a true positive (TP) if it was marked as a spindle by either of the experts and the automated detection algorithm. In this way, we calculate the true negative (TN), false positive (FP) and false negative (FN) values which lead to a 2 by 2 contingency table. These values are then used to evaluate the recall and precision scores of each of the detectors. Since spindle events are rare overnight, the specificity values will be abnormally high and may not lead to a proper analysis of the automated detector. The recall and precision scores are further used to calculate the F_1 score, where the F_1 score is defined as the harmonic mean of recall and precision. Note that the F_1 score ranges from 0 to 1, with 1 indicating a perfect detector. Similar to the F_1 score, we also calculate Cohen's κ [17] and Matthews Correlation Coefficient (MCC) [41].

[6] `http://www.teuniz.net/edfbrowser/`

[7] `https://github.com/aparek/detoks`

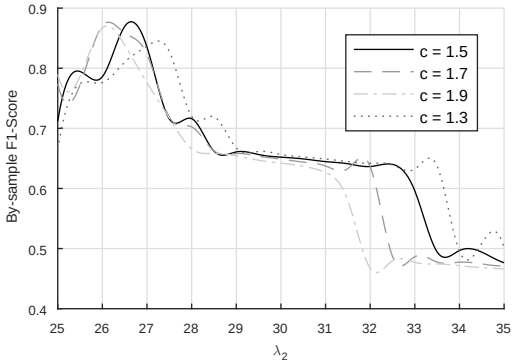

Figure 13: By-sample F_1 score as a function of threshold c and λ_2. For visibility only the c values that yield the highest F_1 scores are shown. The test data is chosen from Excerpt 2 of the online database described in Sec. 5.

5.3. Parameter Tuning

As described in Sec. 4.1 the proposed McSleep spindle detection method requires a set of parameters to be chosen. Recall that the proposed McSleep method is a two-step detection process: first we estimate the transient and the oscillatory component and then use the oscillatory component to detect spindles by using a combination of bandpass filter and the Teager operator. The task-specific parameters such as the passband of the bandpass filter and block length are fixed to the same values as described in Sec. 4.1.

Four key algorithmic-specific parameters are required to be set: λ_0, λ_1, λ_2 and c (threshold for the Teager operator). The parameters λ_0 and λ_1 are set empirically to 0.3 and 6.5 for all the excerpts in the DREAMS database with a sampling of frequency of 200 Hz, whereas they are set as 0.6 and 7 for all the excerpts in the MASS database. For excerpts with a different sampling frequency in the DREAMS database, the values of λ_0 and λ_1 are scaled appropriately. The fixed value of λ_0 and λ_1 is chosen so as to ensure that the oscillatory component is relatively free of transient activity such as BCG or other cardiac artifacts. Note that we run only the transient separation algorithm, and not the entire proposed spindle detection method, in order to fix λ_0 and λ_1.

Once the transient component is estimated the remaining parameters that need to be set are λ_2 and c. For each subject, we choose a segment of the multichannel EEG (usually between 5-10 epochs) and run the proposed McSleep method (transient separation + spindle detection) for a grid of values of λ_2 and c. We select the set of parameters that yield the highest F_1 score for that particular segment. Figure 13 shows the F_1 score as a function of λ_2 and c for a 10 epoch segment for Excerpt 2 from the DREAMS database. It can be seen from Fig. 13 that the optimal set of parameters is $\lambda_2 \approx 26.5$ and $c = 1.5$. For the experiments that follow, the value of λ_2 is varied in the range $(25, 35)$ for the DREAMS database and in the range $(45, 48)$ for the MASS database. The value of the threshold c is varied in the range $(0.5, 3)$ for both

Table 1: Evaluation of proposed method for sleep spindle detection as described in Sec. 5. Average values for the F_1 score Matthews Correlation Coefficient (MCC), recall and precision over 5 excerpts are listed. Standard deviation values are shown in parenthesis.

Methods	By-Sample Performance on DREAMS			
	F_1 score	MCC [41]	Recall	Precision
Wendt [68]	0.49 (0.07)	0.47 (0.06)	0.57 (0.08)	0.44 (0.10)
Martin [40]	0.50 (0.08)	0.50 (0.07)	0.43 (0.11)	0.64 (0.06)
Wamsley [66]	0.06 (0.12)	0.07 (0.16)	0.04 (0.07)	0.18 (0.36)
Bodizs [8]	0.34 (0.15)	0.27 (0.15)	0.75 (0.30)	0.22 (0.10)
Mölle [33]	0.40 (0.26)	0.32 (0.26)	0.28 (0.23)	0.50 (0.33)
Devuyst [24]	0.62 (0.03)	0.60 (0.03)	0.71 (0.07)	0.63 (0.08)
DETOKS [51]	0.70 (0.02)	0.68 (0.02)	0.71 (0.02)	0.68 (0.05)
McSleep	0.66 (0.02)	0.64 (0.02)	0.63 (0.02)	0.69 (0.04)

Table 2: Average values of the F_1 score, Matthews Correlation Coefficient (MCC), recall and precision over 15 EEG recordings are listed. The PSG recordings are from the MASS [46] database for which both the experts annotated spindles on the central channel (C3). (Standard deviation values are listed in parenthesis.)

Methods	By-Sample Performance on MASS			
	F_1 score	MCC [41]	Recall	Precision
Wendt [68]	0.54 (0.06)	0.52 (0.05)	0.55 (0.13)	0.56 (0.12)
Martin [40]	0.32 (0.20)	0.32 (0.21)	0.66 (0.33)	0.29 (0.20)
DETOKS [51]	0.60 (0.06)	0.59 (0.05)	0.55 (0.12)	0.70 (0.11)
McSleep	0.62 (0.06)	0.60 (0.05)	0.61 (0.09)	0.64 (0.08)

the databases. Note that selecting the optimal parameters based on a small segment of the EEG may seem as over-fitting. As such, we suggest running the proposed method on the entire EEG a few times with parameters surrounding the optimal set obtained above.

5.4. Results

The average F_1, MCC, recall and precision values for the proposed McSleep method in comparison with the other existing detectors are listed in Table 1 for the DREAMS database and in Table 2 for the MASS database. Detailed statistical measures for the proposed method in comparison with the existing methods on both the databases are listed in Appendix 9.2.

The proposed detection method on a 30 minute excerpt from the DREAMS database with three EEG channels (sampling frequency of 200 Hz) takes on an average 10 seconds on an Intel Core i7 cpu-based machine. This consists of 40 iterations of the transient separation algorithm, followed by bandpass filtering the oscillatory component and applying the Teager operator for spindle detection. The average runtime for the single channel detectors varied from 10 to 100 seconds with the most time taken by the Mölle detector [33]. For an overnight EEG recording (approx. 8 hours) with six channels the proposed method takes about 4 minutes to detect spindles in all the six channels. In comparison, bandpass-filter-based single-channel detectors take about 1 to 5 minutes for detecting spindles in all 6 channels, while a transient separation based algorithm, such as DETOKS [51], takes on an average 8 minutes (run in parallel over all 30 second epochs of the six channels).

The proposed McSleep detection method can be run in two ways: either in parallel on 30 second epochs or on the entire overnight EEG (aprrox. 8 hours). We choose the former for the analysis presented in this paper. The epoch-by-epoch method of execution is done solely for faster runtimes. Moreover, it also enables the utility of the proposed method in an online mode[8]. The spindle detection is not affected whether the proposed method is run in parallel or on the entire overnight EEG. Furthermore, the proposed method can be run on user-chosen epochs as well.

5.5. Discussion

On the DREAMS database, the proposed McSleep detection method achieved better average F_1 scores compared to other state-of-the-art detectors. The highest average F_1 scores, however, were obtained by the DETOKS [51] method. The Martin [40] and Devuyst [23] methods performed relatively better than other detectors, in terms of the average F_1 scores. On the MASS database, the proposed McSleep method achieved highest F_1 scores. The DETOKS method [51] performed similarly but with slightly lower F_1 scores (0.60 ± 0.06 as compared to 0.62 ± 0.06 for McSleep). However, for PSG 17, from the MASS database, the DETOKS method outperformed the proposed method. Note that the results for only the top four methods are shown in Table 2.

As shown in Sec. 4, the proposed McSleep method performs relatively better by utilizing joint information from multiple EEG channels to separate the transients and oscillations. As a result, the performance of the proposed method is expected to, in terms of spindle detection against other existing detectors, increase proportionally with the number of EEG channels recorded. For the DREAMS database, where only three EEG channels were provided, the proposed method didn't outperform the state-of-the-art single channel detector (DETOKS in this case). On the other hand, for the MASS database, where six recorded EEG channels were used for McSleep, the performance of the proposed method is relatively better in comparison to other detectors. Furthermore, contrary to the belief that using multiple EEG channels might slow down an automated detector, the proposed McSleep method is faster than single channel methods run sequentially on multiple channels.

The transient separation algorithm for the proposed McSleep method and the one proposed in DETOKS [51] are quite similar. In particular, the regularization terms used for the transient component are same in both the methods, with the only difference being that the proposed

[8] An online algorithm is able to process the input data in a piece-by-piece fashion without requiring the presence of an entire signal.

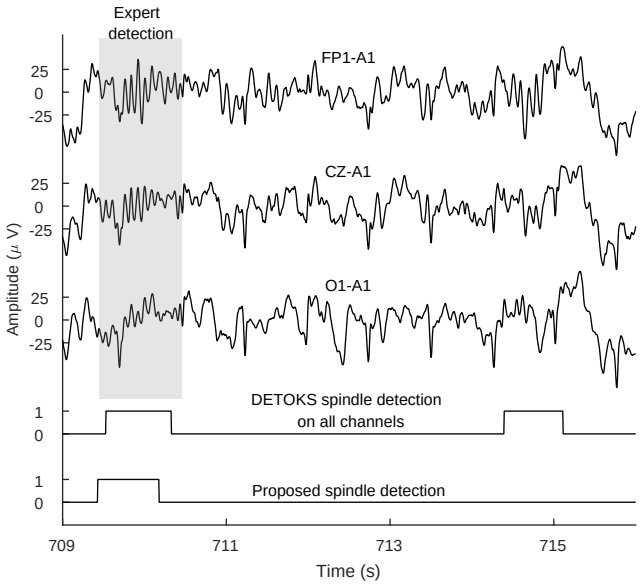

Figure 14: Comparison of the proposed McSleep detection method and the DETOKS detection method. The DETOKS method detects a false positive when run on either a single central channel or on all the channels. The EEG is obtained from excerpt 2 of the online EEG database [24].

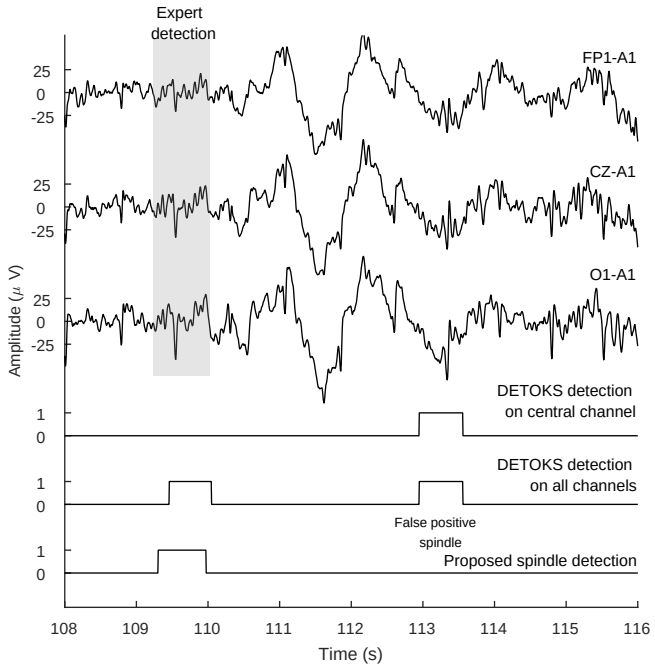

Figure 15: Proposed McSleep detection in comparison with DETOKS method run sequentially on all channels. The proposed detection contains fewer false positive spindles due to better separation of transients.

method uses a multichannel input whereas DETOKS uses a single channel input. The notable difference between the McSleep and DETOKS method is in the use of the regularization term for the oscillatory component with the former using a low-rank regularization and the latter using a sparse STFT regularization. For the case of a single channel input, the proposed objective function in (19) penalizes sum of absolute values of overlapping blocks. In comparison, DETOKS penalizes sum of absolute values of overlapping STFT blocks. As such, the proposed McSleep method and DETOKS are expected to perform similarly, especially for a dataset that either has EEG channels only from one hemisphere of the brain or where the experts viewed only the central channel on the screen while scoring spindles.

A simple method for detecting sleep spindles across all the channels of scalp EEG is to run the existing single channel detectors sequentially channel-by-channel and if required, combine the resulting detections. In order to combine the single channel detections a majority-vote type method may be used. However, ignoring detected spindles on the basis that they do not pass a majority vote can possibly lead to a high type II error, especially for studies that are investigating the spatial distribution of sleep spindles overnight (for e.g., see [9], [52] and the references therein). Another method, perhaps simpler than majority voting, for combining the detection is to use a 'union' rule: a sleep spindle detected in any one channel is a valid detected spindle. However, such a union rule generally leads to high type I error (high false positives are reported).

Consider the EEG segment shown in Fig. 14, where the expert annotated spindle is at 109.3 seconds. Figure 14 shows the spindle detection obtained using three methods: DETOKS run on only the central EEG channel, DETOKS run on all three channels sequentially with separate parameter tuning, and the proposed McSleep method. The single channel DETOKS method run on the central channel alone detects a false positive. Whereas, using the union rule, the all-channel DETOKS detects the true positive spindle but also retains the false positive spindle. In contrast, the proposed method, due to a better separation of transients and oscillations, does not detect the false positive spindle while correctly detecting the expert annotated spindle. Similar behavior is observed in another EEG segment shown in Fig. 15.

6. Conclusion

Automated spindle detectors that consider only a single channel of the recorded EEG are blind to the presence of spindles in other channels. On the other hand, a human expert annotating spindles visually based on the AASM manual may be subliminally influenced by the presence of spindles in other EEG channels visible on the screen. In order to mimic such a human behavior and simultaneously utilize joint information from multiple EEG channels, we propose a multichannel sleep spindle detector using a nonlinear signal model for the multichannel EEG. The proposed spindle method uses a multichannel transient separation algorithm that separates the non-rhythmic transients from spindle-like oscillations, thereby estimating the

components of the proposed non-linear signal model. The transient component is modeled as a piece-wise constant signal with a baseline of zero whereas fixed-length blocks of the multichannel oscillatory component are assumed to be low-rank. The oscillatory component is then used to detect spindles. A fourth order Butterworth bandpass filter and the Teager operator are used to detect spindles following the transient separation process.

Several examples are shown to illustrate the utility of the proposed multichannel spindle detector and a comparison with state-of-the-art single channel spindle detectors is performed using two publicly available online databases. A fast run-time and better average F_1 scores enable the proposed multichannel sleep spindle detector to be a valuable tool for studying the architecture of sleep spindles and tracking their behavior in sleep EEG. While the degree to which a human expert is influenced by the presence of spindles in channels visible to him on the screen is an open question, the proposed method shows that using multichannel EEG certainly yields better estimation of transients and spindles alike. Thereby, increasing the agreement between human expert and an automated algorithm for spindle detection.

7. Acknowledgments

The authors thank the anonymous reviewers for their detailed suggestions which have improved the manuscript considerably.

8. References

[1] N. Acir and C. Güzeli. Automatic recognition of sleep spindles in EEG via radial basis support vector machine based on a modified feature selection algorithm. *Neural Comput. Appl.*, 14(1):56–65, Jan. 2005. ISSN 0941-0643. doi: 10.1007/s00521-004-0442-z.

[2] M. Adamczyk, L. Genzel, M. Dresler, A. Steiger, and E. Friess. Automatic sleep spindle detection and genetic influence estimation using continuous wavelet transform. *Front. Hum. Neurosci.*, 9:624, 2015. ISSN 1662-5161. doi:10.3389/fnhum.2015.00624.

[3] B. Ahmed, A. Redissi, and R. Tafreshi. An automatic sleep spindle detector based on wavelets and the Teager energy operator. In *Int. Conf. IEEE Eng. Med. Biol. Soc.*, pages 2596–2599, Sep. 2009. doi:10.1109/IEMBS.2009.5335331.

[4] T. Andrillon and Yuv. Sleep spindles in humans: Insights from intracranial EEG and unit recordings. *J. Neurosci.*, 31(49):17821–17834, 2011. ISSN 0270-6474. doi:10.1523/jneurosci.2604-11.2011.

[5] B. Babadi, S. M. McKinney, V. Tarokh, and J. M. Ellenbogen. DiBa: A data-driven Bayesian algorithm for sleep spindle detection. *IEEE Trans. Biomed. Eng.*, 59(2):483–493, 2012. doi:10.1109/TBME.2011.2175225.

[6] A. K. Barros, R. Rosipal, M. Girolami, G. Dorffner, and N. Ohnishi. Extraction of sleep-spindles from the electroencephalogram (EEG). In *Artif. Neural Networks Med. Biol.*, Perspectives in Neural Computing, pages 125–130. Springer London, 2000. ISBN 978-1-85233-289-1. doi:10.1007/978-1-4471-0513-8_17.

[7] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.*, 2(1):183–202, 2009. doi:10.1137/080716542.

[8] R. Bódizs, J. Körmendi, P. Rigó, and A. S. Lázár. The individual adjustment method of sleep spindle analysis: methodological improvements and roots in the fingerprint paradigm. *J. Neurosci. Methods*, 178(1):205–13, Mar. 2009. ISSN 0165-0270. doi:10.1016/j.jneumeth.2008.11.006.

[9] R. Bódizs, F. Gombos, P. P. Ujma, S. Szakadát, P. Sándor, P. Simor, A. Pótári, B. N. Konrad, L. Genzel, A. Steiger, and Others. The hemispheric lateralization of sleep spindles in humans. *Sleep Spindl. Cortical Up States*, 1(1):1–13, 2017.

[10] S. Boyd, N. Parikh, E. Chu, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.*, 3(1):1–122, 2010. ISSN 1935-8237. doi:10.1561/2200000016.

[11] A. J. Brockmeier and J. C. Principe. Learning recurrent waveforms within EEGs. *IEEE Trans. Biomed. Eng.*, 63(1):43–54, 2016. ISSN 15582531. doi:10.1109/TBME.2015.2499241.

[12] B. W. Brunton, L. A. Johnson, J. G. Ojemann, and J. N. Kutz. Extracting spatial-temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. *J. Neurosci. Methods*, 258:1–15, 2016. ISSN 1872678X. doi:10.1016/j.jneumeth.2015.10.010.

[13] J.-F. Cai, S. Osher, and Z. Shen. Split Bregman methods and frame based image restoration. *Multiscale Model. Simul.*, 8(2):337–369, Dec. 2010. ISSN 1540-3459. doi:10.1137/090753504.

[14] D. Z. Carvalho, G. J. L. Gerhardt, G. Dellagustin, E. L. de Santa-Helena, N. Lemke, A. Z. Segal, and S. V. Schönwald. Loss of sleep spindle frequency deceleration in obstructive sleep apnea. *Clin. Neurophysiol.*, 125(2):306–312, 2014. ISSN 13882457. doi:10.1016/j.clinph.2013.07.005.

[15] B. C. Clawson, J. Durkin, and S. J. Aton. Form and function of sleep spindles across the lifespan. *Neural Plast.*, (1):1–16, 2016. ISSN 16875443. doi:10.1155/2016/6936381.

[16] Z. Clemens, M. Mölle, L. Eross, P. Barsi, P. Halász, and J. Born. Temporal coupling of parahippocampal ripples, sleep spindles and slow oscillations in humans. *Brain*, 130(11):2868–2878, 2007. ISSN 00068950. doi:10.1093/brain/awm146.

[17] J. Cohen. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.*, 20(1):37–46, 1960. doi:10.1177/001316446002000104.

[18] P. L. Combettes and J. C. Pesquet. A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE J. Sel. Top. Signal Process.*, 1(4):564–574, Dec. 2007. ISSN 19324553. doi:10.1109/JSTSP.2007.910264.

[19] L. Condat. A direct algorithm for 1D total variation denoising. *IEEE Signal Process. Lett.*, 20(11):1054–1057, Nov. 2013. ISSN 1070-9908. doi:10.1109/LSP.2013.2278339.

[20] D. Coppieters, P. Maquet, and C. Phillips. Sleep spindles as an electrographic element: description and automatic detection methods. *Neural Plast.*, 2016:1–19, 2016.

[21] J. Costa, M. Ortigueira, A. Batista, and L. Paiva. An automatic sleep spindle detector based on WT, STFT and WMSD. *World Acad. Sci. Eng. Technol.*, pages 1833–1836, 2012.

[22] R. T. F. De Souza, G. J. L. Gerhardt, S. V. Schönwald, J. L. Rybarczyk-Filho, and N. Lemke. Synchronization and propagation of global sleep spindles. *PLoS One*, 11(3), 2016. ISSN 19326203. doi:10.1371/journal.pone.0151369.

[23] S. Devuyst, T. Dutoit, J. Didier, F. Meers, E. Stanus, P. Stenuit, and M. Kerkhofs. Automatic sleep spindle detection in patients with sleep disorders. *Proc. IEEE Int. Conf. Eng. Med. Biol.*, pages 3883–3886, Aug. 2006. ISSN 1557-170X. doi:10.1109/IEMBS.2006.259298.

[24] S. Devuyst, T. Dutoit, P. Stenuit, and M. Kerkhofs. Automatic K-complexes detection in sleep EEG recordings using likelihood thresholds. In *Proc. IEEE Int. Conf. Eng. Med. Biol.*, pages 4658–4661, 2010. ISBN 9781424441235. doi:10.1109/IEMBS.2010.5626447.

[25] S. Devuyst, T. Dutoit, P. Stenuit, and M. Kerkhofs. Automatic sleep spindles detection - overview and development of a standard proposal assessment method. *Proc. IEEE Int. Conf. Eng. Med. Biol.*, pages 1713–1716, Aug. 2011. ISSN 1557-170X. doi:10.1109/IEMBS.2011.6090491.

[26] D. L. Donoho. De-noising by soft-thresholding. *IEEE Trans. Inf. Theory*, 41(3):613–627, 1995. ISSN 00189448. doi:10.1109/18.382009.

[27] P. Durka and K. Blinowska. Matching pursuit parametrization of sleep spindles. *Eng. Med. Biol.*, 3:1011–1012, Oct. 1996. doi:10.1109/IEMBS.1996.652685.

[28] P. J. Durka, A. Matysiak, E. M. Montes, P. V. Sosa, and K. J. Blinowska. Multichannel matching pursuit and EEG inverse solutions. *J. Neurosci. Methods*, 148(1):49–59, 2005. ISSN 01650270. doi:10.1016/j.jneumeth.2005.04.001.

[29] J. Eckstein and D. P. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.*, 55(1-3):293–318, Apr. 1992. ISSN 00255610. doi:10.1007/BF01581204.

[30] A. Erdamar, F. Duman, and S. Yetkin. A wavelet and teager energy operator based method for automatic detection of K-Complex in sleep EEG. *Expert Syst. Appl.*, 39(1):1284–1290, Jan. 2012. ISSN 09574174. doi:10.1016/j.eswa.2011.07.138. URL `http://dx.doi.org/10.1016/j.eswa.2011.07.138`.

[31] F. Ferrarelli, R. Huber, M. J. Peterson, M. Massimini, M. Murphy, B. a. Riedner, A. Watson, P. Bria, and G. Tononi. Reduced sleep spindle activity in schizophrenia patients. *Am. J. Psychiatry*, 164(3):483–92, Mar. 2007. ISSN 0002-953X. doi:10.1176/appi.ajp.164.3.483.

[32] S. M. Fogel and C. T. Smith. The function of the sleep spindle: a physiological index of intelligence and a mechanism for sleep-dependent memory consolidation. *Neurosci. Biobehav. Rev.*, 35(5):1154–65, Apr. 2011. ISSN 1873-7528. doi:10.1016/j.neubiorev.2010.12.003.

[33] S. Gais, M. Mölle, K. Helms, and J. Born. Learning-dependent increases in sleep spindle density. *J. Neurosci.*, 22(15):6830–4, Aug. 2002. ISSN 1529-2401. doi:20026697.

[34] J. Gilles, T. Meyer, and P. Douglas. Leveraging sparsity: A low-rank + sparse decomposition (LR+SD) method for automatic EEG artifact removal. *Proc. Int. Work. Sparsity Tech. Med. Imaging*, pages 80–88, 2014.

[35] D. Gorur, U. Halici, and H. Aydin. Sleep spindles detection using short time Fourier transform and neural networks. *Int. Jt. Conf. Neural Networks*, pages 1631–1636, 2002.

[36] A. Jaleel, B. Ahmed, R. Tafreshi, D. B. Boivin, L. Streletz, and N. Haddad. Improved spindle detection through intuitive pre-processing of electroencephalogram. *J. Neurosci. Methods*, 233:1–12, Jun. 2014. ISSN 1872-678X. doi:10.1016/j.jneumeth.2014.05.009.

[37] T. Lajnef, S. Chaibi, J.-B. Eichenlaub, P. M. Ruby, P.-E. Aguera, M. Samet, A. Kachouri, and K. Jerbi. Sleep spindle and K-complex detection using tunable Q-factor wavelet transform and morphological component analysis. *Front. Hum. Neurosci.*, 9:1–17, 2015. ISSN 1662-5161. doi:10.3389/fnhum.2015.00414.

[38] T. Lajnef, C. O. Reilly, E. Combrisson, S. Chaibi, A. Kachouri, S. Frenette, J. Carrier, and K. Jerbi. Meet spinky : An open-source spindle and K-complex detection toolbox validated on the open-access Montreal archive of sleep studies (MASS). *Front. Neuroinform.*, 11:1–13, 2017. ISSN 16625196. doi:10.3389/fninf.2017.00015.

[39] D. S. Manoach, J. Q. Pan, S. M. Purcell, and R. Stickgold. Reduced sleep spindles in Schizophrenia: A Treatable Endophenotype That Links Risk Genes to Impaired Cognition? *Biol. Psychiatry*, 80(8):599–608, 2016. ISSN 18732402. doi:10.1016/j.biopsych.2015.10.003.

[40] N. Martin, M. Lafortune, J. Godbout, M. Barakat, R. Robillard, G. Poirier, C. Bastien, and J. Carrier. Topography of age-related changes in sleep spindles. *Neurobiol. Aging*, 34(2):468–476, Feb. 2013. ISSN 1558-1497. doi:10.1016/j.neurobiolaging.2012.05.020.

[41] B. W. Matthews. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta - Protein Struct.*, 405(2):442–451, 1975. doi:10.1016/0005-2795(75)90109-9.

[42] S. Noachtar, C. Binnie, J. Ebersole, F. Mauguiere, A. Sakamoto, and B. Westmoreland. A glossary of terms most commonly used by clinical electroencephalographers and proposal for the report form for the EEG findings. The International Federation of Clinical Neurophysiology. *Electroencephalogr. Clin. Neurophysiol. Suppl.*, 52:21–41, 1999. ISSN 0424-8155 (Print).

[43] A. Nonclercq, C. Urbain, D. Verheulpen, C. Decaestecker, P. Van Bogaert, and P. Peigneux. Sleep spindle detection through amplitude-frequency normal modelling. *J. Neurosci. Methods*, 214(2):192–203, 2013. ISSN 01650270. doi:10.1016/j.jneumeth.2013.01.015.

[44] C. O'Reilly and T. Nielsen. Assessing EEG sleep spindle propagation. Part 1: Theory and proposed methodology. *J. Neurosci. Methods*, 221:202–214, 2014. ISSN 1872678X. doi:10.1016/j.jneumeth.2013.08.013.

[45] C. O'Reilly and T. Nielsen. Automatic sleep spindle detection: benchmarking with fine temporal resolution using open science tools. *Front. Hum. Neurosci.*, 9:353, 2015. ISSN 1662-5161. doi:10.3389/fnhum.2015.00353.

[46] C. O'Reilly, N. Gosselin, J. Carrier, and T. Nielsen. Montreal archive of sleep studies: An open-access resource for instrument benchmarking and exploratory research. *J. Sleep Res.*, 23(6):628–635, 2014. ISSN 13652869. doi:10.1111/jsr.12169.

[47] C. O'Reilly, J. Godbout, J. Carrier, and J.-M. Lina. Combining time-frequency and spatial information for the detection of sleep spindles. *Front. Hum. Neurosci.*, 9:70, 2015. ISSN 1662-5161. doi:10.3389/fnhum.2015.00070.

[48] A. Parekh and I. W. Selesnick. Convex fused lasso denoising with non-convex regularization and its use for pulse detection. In *IEEE Symp. Signal Process. Med. Biol.*, pages 1–6, 2015.

[49] A. Parekh and I. W. Selesnick. Enhanced low-rank matrix approximation. *IEEE Signal Process. Lett.*, 23(4):493–497, 2016.

[50] A. Parekh, I. W. Selesnick, D. M. Rapoport, and I. Ayappa. Sleep spindle detection using time-frequency sparsity. In *Proc. IEEE Signal Process. Med. Biol. Symp.*, Dec. 2014. ISBN 9781479981847. doi:10.1109/SPMB.2014.7002965.

[51] A. Parekh, I. W. Selesnick, D. M. Rapoport, and I. Ayappa. Detection of K-complexes and sleep spindles (DETOKS) using sparse optimization. *J. Neurosci. Methods*, 251:37–46, 2015. ISSN 1872678X. doi:10.1016/j.jneumeth.2015.04.006.

[52] S. M. Purcell, D. S. Manoach, C. Demanuele, B. Cade, S. Mariani, R. Cox, R. Saxena, J. Pan, J. Smoller, S. Redline, and R. Stickgold. Characterizing sleep spindles in 11,630 individuals from the National Sleep Research Resource. *Preprint*, 2016.

[53] L. Ray, S. Sockeel, M. Soon, A. Bore, A. Myhr, B. Stojanoski, R. Cusack, A. M. Owen, J. Doyon, and S. M. Fogel. Expert and crowd-sourced validation of an individualized sleep spindle detection method employing complex demodulation and individualized normalization. *Front. Hum. Neurosci.*, 9(507):1–16, 2015. ISSN 1662-5161. doi:10.3389/fnhum.2015.00507.

[54] L. B. Ray, S. M. Fogel, C. T. Smith, and K. R. Peters. Validating an automated sleep spindle detection algorithm using an individualized approach. *J. Sleep Res.*, 19(2):374–8, Jun. 2010. ISSN 1365-2869. doi:10.1111/j.1365-2869.2009.00802.x.

[55] R. Rosipal, G. Dorffner, and E. Trenker. Can ICA improve sleep-spindles detection? *Neural Netw. World*, 8(5):539–547, 1998. ISSN 12100552.

[56] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D Nonlinear Phenom.*, 60(1-4):259–268, Nov. 1992. ISSN 01672789. doi:10.1016/0167-2789(92)90242-F.

[57] P. Schimicek, J. Zeitlhofer, P. Anderer, and B. Saletu. Automatic sleep-spindle detection procedure: aspects of reliability and validity. *Clin. Electroencephalogr.*, 25:26–29, 1994. ISSN 1550-0594. doi:10.1177/155005949402500108.

[58] J. Shi, X. Liu, Y. Li, Q. Zhang, Y. Li, and S. Ying. Multichannel EEG-based sleep stage classification with joint collaborative representation and multiple kernel learning. *J. Neurosci. Methods*, 254:94–101, 2015. ISSN 1872678X. doi:10.1016/j.jneumeth.2015.07.006.

[59] M. H. Silber, S. Ancoli-Israel, M. H. Bonnet, S. Chokroverty, M. M. Grigg-Damberger, M. Hirshkowitz, S. Kapen, S. a. Keenan, M. H. Kryger, T. Penzel, M. R. Pressman, and C. Iber. The visual scoring of sleep in adults. *J. Clin. Sleep Med.*, 3(2):121–31, Mar. 2007. ISSN 1550-9389.

[60] J. Starck, M. Elad, and D. Donoho. Image decomposition via the combination of sparse representations and a variational approach. *IEEE Trans. Image Process.*, 14(10):1570–82, Oct. 2005. ISSN 1057-7149.

[61] C. Stepnowsky, D. Levendowski, D. Popovic, I. Ayappa, and D. M. Rapoport. Scoring accuracy of automated sleep staging from a bipolar electroocular recording compared to manual scoring by multiple raters. *Sleep Med.*, 14(11):1199–207, Nov. 2013. ISSN 1878-5506. doi:10.1016/j.sleep.2013.04.022.

[62] D. Studer, U. Hoffmann, and T. Koenig. From EEG dependency multichannel matching pursuit to sparse topographic EEG decomposition. *J. Neurosci. Methods*, 153(2):261–275, 2006. ISSN 01650270. doi:10.1016/j.jneumeth.2005.11.006.

[63] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, and K. Knight. Sparsity and smoothness via the fused lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 67(1):91–108, 2005. ISSN 13697412. doi:10.1111/j.1467-9868.2005.00490.x.

[64] A. Tsanas and G. D. Clifford. Stage-independent, single lead EEG sleep spindle detection using the continuous wavelet transform and local weighted smoothing. *Front. Hum. Neurosci.*, 9:181, 2015. ISSN 1662-5161. doi:10.3389/fnhum.2015.00181.

[65] D. Tylavsky and G. Sohie. Generalization of the matrix inversion lemma. *Proc. IEEE*, 74:1050–1052, 1986. ISSN 0018-9219. doi:10.1109/PROC.1986.13587.

[66] E. Wamsley, M. Tucker, and A. Shinn. Reduced sleep spindles and spindle coherence in schizophrenia: mechanisms of impaired memory consolidation? *Biol. Psychiatry*, 71(2):154–161, 2012. doi:10.1016/j.biopsych.2011.08.008.Reduced.

[67] S. C. Warby, S. L. Wendt, P. Welinder, E. G. S. Munk, O. Carrillo, H. B. D. Sorensen, P. Jennum, P. E. Peppard, P. Perona, and E. Mignot. Sleep-spindle detection: crowdsourcing and evaluating performance of experts, non-experts and automated methods. *Nat. Methods*, 11(4):385–92, Apr. 2014. ISSN 1548-7105. doi:10.1038/nmeth.2855.

[68] S. L. Wendt, J. E. Christensen, J. Kempfner, H. L. Leonthin, P. Jennum, and H. B. D. Sorensen. Validation of a novel automatic sleep spindle detector with high performance during sleep in middle aged subjects. *Proc. IEEE Int. Conf. Eng. Med. Biol.*, pages 4250–4253, Aug. 2012. ISSN 1557-170X. doi:10.1109/EMBC.2012.6346905.

[69] M. Wohlleber, A. Parekh, T. Gumb, A. Varga, and R. Osorio. Increased Cerebrospinal P-Tau is associated with reduced spindle frequency in cognitively normal elderly. In *Int. Conf. Sleep Spindl.*, page 1, Budapest, May 2016.

[70] M. Woodbury. Inverting modified matrices. In *Memo. Rep.*, volume 42, page 4. Princeton Unversity, 1950.

[71] M. Younes, J. Raneri, and P. Hanly. Staging sleep in polysomnograms: Analysis of inter-scorer variability. *J. Clin. Sleep Med.*, 12(6):885–894, 2016. ISSN 15509397. doi:10.5664/jcsm.5894.

[72] M. Younes. The case for using digital EEG analysis in clinical sleep medicine. *Sleep Science and Practice*, 1(2):1–15, 2017. ISSN 2398-2683. doi:10.1186/s41606-016-0005-0.

9. Appendix

9.1. Solution to the Least-Squares step of the proposed transient separation algorithm

We derive the solution to the least-squares sub-problem in (22a), which is written below for clarity to the reader.

$$U,V \leftarrow \arg\min_{U,V}\Bigg\{\frac{1}{2}\|Y-(U+H(V))\|_2^2 + \frac{\mu}{2}\|U-X-D_1\|_2^2+\frac{\mu}{2}\|V-C-D_2\|_2^2\Bigg\}.$$

We make the following substitutions

$$\hat{U}=[U,V]^T, \qquad \hat{D}=[D_1,D_2]^T, \tag{31a}$$

$$\hat{X}=[X,C]^T, \qquad M=[I,H], \tag{31b}$$

and re-write the objective function as

$$\hat{U} \leftarrow \arg\min_{\hat{U}}\Bigg\{\frac{1}{2}\|Y-M\hat{U}\|_2^2+\frac{\mu}{2}\|\hat{U}-\hat{X}-\hat{D}\|_2^2\Bigg\}. \tag{32}$$

The solution to (32) can be written explicitly as

$$\hat{U}=\left[M^TM+\mu I\right]^{-1}\left[M^TY+\mu(\hat{X}+\hat{D})\right]. \tag{33}$$

The inverse in (33) results in inverting a dense matrix consisting of M^TM. In order to efficiently compute the explicit solution, we use the Matrix Inverse lemma [70, 65]. As such, the solution in (33) can be written as

$$\left[M^TM+\mu I\right]^{-1}=\frac{1}{\mu}\left(I-M^T\left(\mu I+MM^T\right)^{-1}M\right). \tag{34}$$

Note that the operator H, where $H=\Phi^T$, is implemented in this paper for perfect reconstruction. Hence, we have

$$MM^T=[I \quad H]\begin{bmatrix} I \\ H^T \end{bmatrix} \tag{35}$$

$$=2I. \tag{36}$$

As a result, the inverse in (33) can be written as

$$\left[M^TM+\mu I\right]^{-1}=\frac{1}{\mu}\left(I-\frac{1}{\mu+2}M^TM\right), \tag{37}$$

which leads to the following explicit solution for $\hat{U}$,

$$\hat{U}=\frac{1}{\mu}M^TY+\hat{X}+\hat{D} - \frac{1}{\mu+2}\Big(M^TY+\mu M^TM(\hat{X}+\hat{D})\Big). \tag{38}$$

Combining (38) and (31), we get the following steps for

obtaining the solution to the objective function in (22a).

$$f_1 = \frac{1}{\mu} Y + X + D_1 \tag{39a}$$

$$f_2 = \frac{1}{\mu} H^T Y + C + D_2 \tag{39b}$$

$$U = f_1 - \frac{1}{\mu + 2}\left(f_1 + H f_2\right) \tag{39c}$$

$$V = f_2 - \frac{1}{\mu + 2} H^T \left(f_1 + H f_2\right) \tag{39d}$$

Note that $H^T Y$ can be pre-computed outside the iterative loop for speed.

9.2. Illustration of McSleep for sleep spindle detection

The tables in Fig. 16 and Fig. 17 illustrate the performance of the proposed multichannel spindle detection (McSleep) on two online databases: the DREAMS database [23] and the MASS dataset [46] respectively. Detailed statistical measures are also shown for several state-of-the-art spindle detectors in comparison with McSleep. For a summary of each of the measures of performance, we refer the reader to [67].

Excerpt 1	Wendt	Martin	Wamsley	Bodizs	Molle	Devuyst	DETOKS	McSleep
TP	8357	4121	2374	9550	4991	9692	9354	8548
TN	156774	164771	166242	137540	164911	159737	162632	162826
FP	9751	1754	283	28985	1614	6788	3892	3699
FN	5118	9354	11101	3925	8484	3783	4121	4927
Recall	0.62	0.31	0.18	0.71	0.37	0.72	0.69	0.63
Precision	0.46	0.70	0.89	0.25	0.76	0.59	0.71	0.70
F1 Score	**0.53**	**0.43**	**0.29**	**0.37**	**0.50**	**0.65**	**0.70**	**0.66**
Specificity	0.94	0.99	1.00	0.83	0.99	0.96	0.98	0.98
NPV	0.97	0.95	0.94	0.97	0.95	0.98	0.98	0.97
Accuracy	0.92	0.94	0.94	0.82	0.94	0.94	0.96	0.95
Kappa	0.49	0.40	0.28	0.29	0.47	0.62	0.68	0.64
MCC	0.49	0.44	0.38	0.34	0.51	0.62	0.68	0.64

Excerpt 2	Wendt	Martin	Wamsley	Bodizs	Molle	Devuyst	DETOKS	McSleep
TP	9881	7954	0	11045	7006	9862	10307	9569
TN	320771	339807	344818	268002	340517	336140	339091	341242
FP	24754	5718	707	77523	5008	9385	6433	4283
FN	4594	6521	14475	3430	7469	4613	4168	4906
Recall	0.68	0.55	0.00	0.76	0.48	0.68	0.71	0.66
Precision	0.29	0.58	0.00	0.13	0.58	0.51	0.62	0.69
F1 Score	**0.40**	**0.57**	**0.00**	**0.21**	**0.53**	**0.59**	**0.67**	**0.68**
Specificity	0.93	0.98	1.00	0.78	0.99	0.97	0.98	0.99
NPV	0.99	0.98	0.96	0.99	0.98	0.99	0.99	0.99
Accuracy	0.92	0.97	0.96	0.78	0.97	0.96	0.97	0.97
Kappa	0.37	0.55	0.00	0.16	0.51	0.57	0.65	0.66
MCC	0.41	0.55	-0.01	0.25	0.51	0.57	0.65	0.66

Excerpt 3	Wendt	Martin	Wamsley	Bodizs	Molle	Devuyst	DETOKS	McSleep
TP	1005	680	12	0	0	1400	1707	1380
TN	86050	87234	86741	0	87717	86882	86772	86976
FP	1667	483	976	0	0	835	944	741
FN	1278	1603	2271	0	2283	883	576	903
Recall	0.44	0.30	0.01	0.00	0.00	0.61	0.75	0.60
Precision	0.38	0.59	0.01	0.00	0.00	0.63	0.64	0.65
F1 Score	**0.41**	**0.40**	**0.01**	**0.00**	**0.00**	**0.62**	**0.69**	**0.63**
Specificity	0.98	0.99	0.99	0.00	1.00	0.99	0.99	0.99
NPV	0.99	0.98	0.97	0.00	0.98	0.99	0.99	0.99
Accuracy	0.97	0.98	0.96	0.00	0.98	0.98	0.98	0.98
Kappa	0.39	0.38	-0.01	0.00	0.00	0.61	0.68	0.62
MCC	0.39	0.41	-0.01	0.00	0.00	0.61	0.69	0.62

Excerpt 5	Wendt	Martin	Wamsley	Bodizs	Molle	Devuyst	DETOKS	McSleep
TP	10461	9672	0	14779	0	10332	14164	13100
TN	330466	333514	337779	295704	340039	335566	333646	332886
FP	9573	6525	2260	44335	0	4473	6392	7153
FN	9500	10289	19961	5182	19961	9629	5797	6861
Recall	0.52	0.49	0.00	0.74	0.00	0.52	0.71	0.66
Precision	0.52	0.60	0.00	0.25	0.00	0.70	0.69	0.65
F1 Score	**0.52**	**0.54**	**0.00**	**0.37**	**0.00**	**0.59**	**0.70**	**0.65**
Specificity	0.97	0.98	0.99	0.87	1.00	0.99	0.98	0.98
NPV	0.97	0.97	0.94	0.98	0.95	0.97	0.98	0.98
Accuracy	0.95	0.95	0.94	0.86	0.95	0.96	0.97	0.96
Kappa	0.50	0.51	-0.01	0.32	0.00	0.57	0.68	0.63
MCC	0.50	0.51	-0.02	0.38	0.00	0.58	0.68	0.63

Excerpt 6	Wendt	Martin	Wamsley	Bodizs	Molle	Devuyst	DETOKS	McSleep
TP	12810	11724	0	17988	11928	13381	15435	13647
TN	327975	332715	336704	288379	331078	332761	332467	333002
FP	9588	4848	859	49184	6485	4802	5095	4561
FN	9627	10713	22437	4449	10509	9056	7002	8790
Recall	0.57	0.52	0.00	0.80	0.53	0.60	0.69	0.61
Precision	0.57	0.71	0.00	0.27	0.65	0.74	0.75	0.75
F1 Score	**0.57**	**0.60**	**0.00**	**0.40**	**0.58**	**0.66**	**0.72**	**0.67**
Specificity	0.97	0.99	1.00	0.85	0.98	0.99	0.99	0.99
NPV	0.97	0.97	0.94	0.99	0.97	0.97	0.98	0.97
Accuracy	0.95	0.96	0.94	0.85	0.95	0.96	0.97	0.96
Kappa	0.54	0.58	-0.01	0.34	0.56	0.64	0.70	0.65
MCC	0.54	0.59	-0.01	0.41	0.56	0.64	0.70	0.66

Figure 16: Performance of proposed McSleep and existing sleep spindle detectors on the DREAMS database.

PSG 1	Wendt	Martin	DETOKS	McSleep
TP	336143	697076	426320	477071
TN	6531970	5256421	6575805	6500153
FP	132700	1408249	88865	164517
FN	412435	51502	322258	271507
Recall	0.45	0.93	0.57	0.64
Precision	0.72	0.33	0.83	0.74
F1 Score	**0.55**	**0.49**	**0.67**	**0.69**
Specificity	0.98	0.79	0.99	0.98
NPV	0.94	0.99	0.95	0.96
Accuracy	0.93	0.80	0.94	0.94
Kappa	0.51	0.40	0.65	0.65
MCC	0.53	0.48	0.66	0.66

PSG 2	Wendt	Martin	DETOKS	McSleep
TP	349918	19949	383521	440165
TN	8136782	8258907	8191375	8122264
FP	128736	6611	74143	143254
FN	349684	679653	316081	259437
Recall	0.50	0.03	0.55	0.63
Precision	0.73	0.75	0.84	0.75
F1 Score	**0.59**	**0.05**	**0.66**	**0.69**
Specificity	0.98	1.00	0.99	0.98
NPV	0.96	0.92	0.96	0.97
Accuracy	0.95	0.92	0.96	0.96
Kappa	0.57	0.05	0.64	0.66
MCC	0.58	0.14	0.66	0.67

PSG 3	Wendt	Martin	DETOKS	McSleep
TP	54648	143117	44057	54865
TN	9205459	7922651	9254322	9241275
FP	55785	1338593	6922	19968
FN	95692	7223	106283	95475
Recall	0.36	0.95	0.29	0.36
Precision	0.49	0.10	0.86	0.73
F1 Score	**0.42**	**0.18**	**0.44**	**0.49**
Specificity	0.99	0.86	1.00	1.00
NPV	0.99	1.00	0.99	0.99
Accuracy	0.98	0.86	0.99	0.99
Kappa	0.41	0.15	0.43	0.48
MCC	0.42	0.28	0.50	0.51

PSG 5	Wendt	Martin	DETOKS	McSleep
TP	130306	256177	147155	206022
TN	7626378	7263744	7655699	7563490
FP	87599	450233	58278	150487
FN	154693	28822	137844	78977
Recall	0.46	0.90	0.52	0.72
Precision	0.60	0.36	0.72	0.58
F1 Score	**0.52**	**0.52**	**0.60**	**0.64**
Specificity	0.99	0.94	0.99	0.98
NPV	0.98	1.00	0.98	0.99
Accuracy	0.97	0.94	0.98	0.97
Kappa	0.50	0.49	0.59	0.63
MCC	0.51	0.55	0.60	0.63

PSG 6	Wendt	Martin	DETOKS	McSleep
TP	77273	191933	72766	112748
TN	7137691	5901258	7193296	7093949
FP	78934	1315367	23329	122676
FN	128054	13394	132561	92579
Recall	0.38	0.93	0.35	0.55
Precision	0.49	0.13	0.76	0.48
F1 Score	**0.43**	**0.22**	**0.48**	**0.51**
Specificity	0.99	0.82	1.00	0.98
NPV	0.98	1.00	0.98	0.99
Accuracy	0.97	0.82	0.98	0.97
Kappa	0.41	0.18	0.47	0.50
MCC	0.42	0.31	0.51	0.50

PSG 7	Wendt	Martin	DETOKS	McSleep
TP	363991	429370	324554	361003
TN	6325675	5100780	6506597	6433853
FP	349695	1574590	168773	241517
FN	206463	141084	245900	209451
Recall	0.64	0.75	0.57	0.63
Precision	0.51	0.21	0.66	0.60
F1 Score	**0.57**	**0.33**	**0.61**	**0.62**
Specificity	0.95	0.76	0.97	0.96
NPV	0.97	0.97	0.96	0.97
Accuracy	0.92	0.76	0.94	0.94
Kappa	0.53	0.24	0.58	0.58
MCC	0.53	0.31	0.58	0.58

PSG 9	Wendt	Martin	DETOKS	McSleep
TP	286407	458648	296587	346870
TN	6939524	5581468	6999632	6885640
FP	120841	1478897	60733	174725
FN	291244	119003	281064	230781
Recall	0.50	0.79	0.51	0.60
Precision	0.70	0.24	0.83	0.67
F1 Score	**0.58**	**0.36**	**0.63**	**0.63**
Specificity	0.98	0.79	0.99	0.98
NPV	0.96	0.98	0.96	0.97
Accuracy	0.95	0.79	0.96	0.95
Kappa	0.55	0.28	0.61	0.60
MCC	0.56	0.36	0.63	0.60

PSG 10	Wendt	Martin	DETOKS	McSleep
TP	343500	402160	293934	332637
TN	5823256	5829748	5927394	5879256
FP	211797	205305	107659	155797
FN	260039	201379	309605	270902
Recall	0.57	0.67	0.49	0.55
Precision	0.62	0.66	0.73	0.68
F1 Score	**0.59**	**0.66**	**0.58**	**0.61**
Specificity	0.96	0.97	0.98	0.97
NPV	0.96	0.97	0.95	0.96
Accuracy	0.93	0.94	0.94	0.94
Kappa	0.55	0.63	0.55	0.57
MCC	0.55	0.63	0.57	0.58

PSG 11	Wendt	Martin	DETOKS	McSleep
TP	303891	458492	344817	377297
TN	5864315	5352844	5831530	5811734
FP	145462	656933	178247	198042
FN	197948	43347	157022	124542
Recall	0.61	0.91	0.69	0.75
Precision	0.68	0.41	0.66	0.66
F1 Score	**0.64**	**0.57**	**0.67**	**0.70**
Specificity	0.98	0.89	0.97	0.97
NPV	0.97	0.99	0.97	0.98
Accuracy	0.95	0.89	0.95	0.95
Kappa	0.61	0.52	0.64	0.67
MCC	0.61	0.57	0.65	0.68

PSG 12	Wendt	Martin	DETOKS	McSleep
TP	286312	326248	228350	222662
TN	7160940	7130682	7520823	7546189
FP	518464	548722	158581	133215
FN	106988	67052	164950	170638
Recall	0.73	0.83	0.58	0.57
Precision	0.36	0.37	0.59	0.63
F1 Score	**0.48**	**0.51**	**0.59**	**0.59**
Specificity	0.93	0.93	0.98	0.98
NPV	0.99	0.99	0.98	0.98
Accuracy	0.92	0.92	0.96	0.96
Kappa	0.44	0.48	0.56	0.57
MCC	0.47	0.52	0.56	0.58

PSG 13	Wendt	Martin	DETOKS	McSleep
TP	405701	493969	368787	367195
TN	8188337	7715593	8428979	8492614
FP	531182	1003926	290540	226905
FN	187036	98768	223950	225542
Recall	0.68	0.83	0.62	0.62
Precision	0.43	0.33	0.56	0.62
F1 Score	**0.53**	**0.47**	**0.59**	**0.62**
Specificity	0.94	0.88	0.97	0.97
NPV	0.98	0.99	0.97	0.97
Accuracy	0.92	0.88	0.94	0.95
Kappa	0.49	0.42	0.56	0.59
MCC	0.51	0.48	0.56	0.59

PSG 14	Wendt	Martin	DETOKS	McSleep
TP	394054	263526	420115	343085
TN	6102700	4108316	6242061	6496306
FP	580375	2574759	441013	186769
FN	136439	266967	110378	187408
Recall	0.74	0.50	0.79	0.65
Precision	0.40	0.09	0.49	0.65
F1 Score	**0.52**	**0.16**	**0.60**	**0.65**
Specificity	0.91	0.61	0.93	0.97
NPV	0.98	0.94	0.98	0.97
Accuracy	0.90	0.61	0.92	0.95
Kappa	0.47	0.04	0.56	0.62
MCC	0.50	0.06	0.58	0.62

PSG 17	Wendt	Martin	DETOKS	McSleep
TP	192741	8300	219858	273181
TN	6704554	6816805	6757313	6569975
FP	143911	31660	91151	278490
FN	214858	399299	187741	134418
Recall	0.47	0.02	0.54	0.67
Precision	0.57	0.21	0.71	0.50
F1 Score	**0.52**	**0.04**	**0.61**	**0.57**
Specificity	0.98	1.00	0.99	0.96
NPV	0.97	0.94	0.97	0.98
Accuracy	0.95	0.94	0.96	0.94
Kappa	0.49	0.03	0.59	0.54
MCC	0.49	0.05	0.60	0.55

PSG 18	Wendt	Martin	DETOKS	McSleep
TP	343539	50317	293794	289144
TN	6669282	4973960	6890077	6968436
FP	433731	2129053	212936	134577
FN	119784	413006	169529	174179
Recall	0.74	0.11	0.63	0.62
Precision	0.44	0.02	0.58	0.68
F1 Score	**0.55**	**0.04**	**0.61**	**0.65**
Specificity	0.94	0.70	0.97	0.98
NPV	0.98	0.92	0.98	0.98
Accuracy	0.93	0.66	0.95	0.96
Kappa	0.52	-0.07	0.58	0.63
MCC	0.54	-0.10	0.58	0.63

PSG 19	Wendt	Martin	DETOKS	McSleep
TP	132165	223550	165437	190335
TN	6402943	4340528	6399020	6348931
FP	63590	2126005	67513	117602
FN	158166	66781	124894	99996
Recall	0.46	0.77	0.57	0.66
Precision	0.68	0.10	0.71	0.62
F1 Score	**0.54**	**0.17**	**0.63**	**0.64**
Specificity	0.99	0.67	0.99	0.98
NPV	0.98	0.98	0.98	0.98
Accuracy	0.97	0.68	0.97	0.97
Kappa	0.53	0.10	0.62	0.62
MCC	0.54	0.19	0.62	0.62

Figure 17: Performance of proposed McSleep method and existing sleep spindle detectors on the MASS database. (The numbers are best viewed on the pdf file by zooming in.)