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 Mathematics, Tandon School of Engineering, New York University*
[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: We consider the detection of sleep spindles simultaneously across the frontal, central and occipital channels of sleep EEG in a single run.
New Method: We propose a multichannel sleep spindle detection method utilizing a multichannel transient separation algorithm based on a sparse optimization framework. The proposed transient separation algorithm decomposes the multichannel EEG into the sum of an oscillatory and a transient component. Consecutive overlapping blocks of the multichannel oscillatory component are assumed to be of low-rank whereas the transient component is assumed to be piecewise constant with a zero baseline. We estimate both the components by minimizing a convex objective function using an iterative algorithm. The multichannel oscillatory component is used in conjunction with the Teager operator for detecting sleep spindles.
Results and comparison with other methods: The performance of the proposed method is evaluated using an online single channel EEG database and compared with 7 state-of-the-art automated detectors. The by-event F_1 scores for the proposed spindle detection method averaged 0.67 ± 0.03. The average false discovery rate for the proposed method was $31.3 \pm 0.04\%$. For an overnight 6 to 8 channel EEG signal, the proposed algorithm takes on an average 2 minutes to detect sleep spindles.
Conclusions: Comparable F_1 scores and fast run times make the proposed spindle detector a valuable tool in answering the open question of studying the dynamics of sleep spindles and tracking their propagation overnight across the scalp in sleep EEG.

Keywords: Sleep 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 [56]. 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 [64]. Sleep spindles reflect a heritable set of traits which is implicated in both sleep regulation and normal cognitive functioning [38]. Recent studies have linked spindle density (number of spindles per minute), duration and amplitude of spindles, and peak frequency of spindles to memory consolidation during sleep [32, 15], cognition in schizophrenia patients [38, 63], brain dysfunction in obstructive sleep apnea [14] and biomarkers for Alzheimer's disease [66]. As a result, understanding the characteristics of sleep spindles is a key in studying their relation to several neuropsychiatric diseases.

Traditionally, sleep spindles are detected in clinics using visual heuristics: number of peaks or bumps of the EEG signal are counted within a specified time window. This process is not only subjective, but also prone to errors and time-consuming. Moreover, visual inspection underscores the fine details of putated spindles [50]. It is not uncommon for studies to utilize more than one expert for detecting spindles. However, the high variability of inter-scorer agreement adds to the complexity of detecting spindles: Cohen's κ coefficient for manual scoring usually ranges between 0.46 and 0.89 [58, 42]. The presence of reliable automated spindle detectors may not only reduce the scoring variability [68] but may also aid in complex longitudinal studies that involve studying global or local sleep spindle dynamics [50, 22, 43].

1.1. Single channel sleep spindle detectors

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 infor-

*Corresponding author. Email address: ankit.parekh@nyu.edu
Source Code available at `https://github.com/aparek/mcsleep.git`

mation 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 artifact removal [36]. One of the first automated detectors to be proposed used a bandpass filter in conjunction with an amplitude threshold [54]. This idea is still the basis of a majority of the bandpass filtering-based automated detectors [65, 23, 39, 31, 16, 33]. Advanced methods utilizing time-frequency information either use a wavelet transform [37, 2, 4, 30, 61, 3] or a short-time Fourier transform (STFT) [21, 45, 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 [47, 48, 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 [57] of the transients and spindles to overcome the drawbacks of filtering and Fast Fourier Transform (FFT) based techniques [51]. 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].

1.2. Multichannel sleep spindle detectors

Recent studies are increasingly looking at the global spindle activity across all channels of scalp EEG in a 10-20 system as a measure to characterize individual sleep spindle density [50, 9]. Automated spindle detection across all the channels can not only help reduce the inter-rater variability among the experts [20] but may also aid in complex studies involving spindle dynamics.

Although, single channel spindle detectors may be used for studying global spindle activity and tracking their propagation overnight, their usage comes at a cost. When detecting spindles on a channel-by-channel basis using an automated algorithm two key issues arise: combining the resulting detection of spindles from all the channels and parameter-tuning for each channel. While it is true that a voting rule can be formulated to combine the spindles detections across the channels, it adds an additional degree of freedom for automated detectors. Moreover, such a voting rule must also be used for expert detection for a fair comparison. Further, 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. Since sleep spindles may occur in different channels or groups of channels at different times [12], analyzing spindle networks (global or local) overnight is challenging with single channel spindle detection methods.

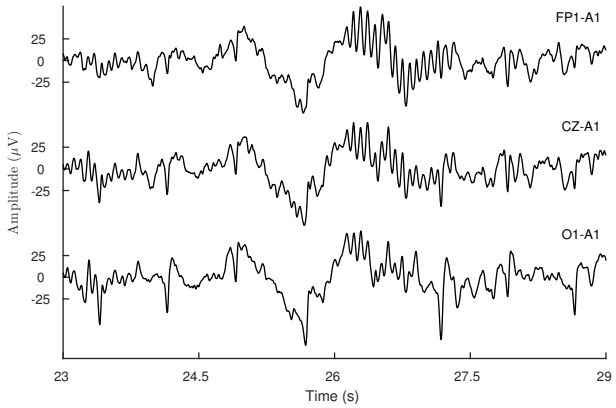

Figure 1: A multichannel EEG with one frontal, one central and one occipital channel. Expert annotated sleep spindle at 26 seconds has different amplitude in different channels.

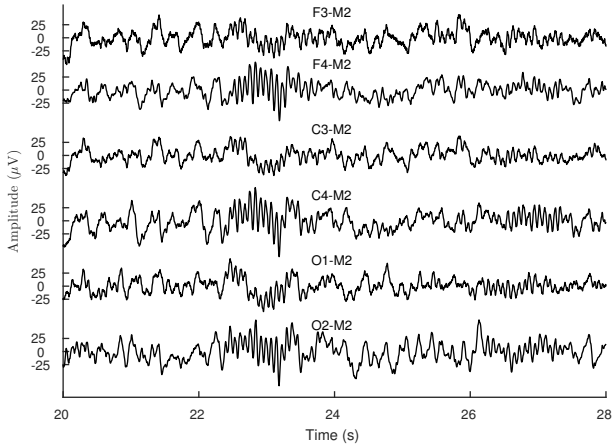

Figure 2: A sample multichannel EEG with two frontal, central and occipital channels. Sleep spindle amplitude is different in left (F3,C3, O1) and right (F4,C4, O2) channels. The EEG excerpt shown here is not from the same dataset used for Fig. 1.

Consider the 6-channel EEG shown in Fig. 1 and Fig. 2 respectively. Classifying spindles as either global or local [15] 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 [64]. In fact, most detectors utilize at the most two channels [44, 51]. Specifically, for the EEG shown in Fig. 2 the spindle at about 26 seconds is prominent only in C4 channel. The particular spindle does appear in other channels but with significantly lower amplitude. Hence, without careful parameter tuning an automated detector may not be able to properly detect the spindles.

Several studies have pointed to the benefit of separating the transients and oscillations sleep spindle detection [48, 47, 20], though few have advocated the use of multichannel EEG [6]. An ICA based approach was studied for the detection of spindles from multichannel EEG in [52].

For sleep-staging and classification a machine learning approach utilizing multichannel EEG was proposed in [55]. 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 [59]. 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].

1.3. Contribution

In this paper we propose a multichannel spindle detection method that detects sleep spindles across all the channels of human scalp EEG in a single run. To this end, we propose a non-linear signal model for the multichannel EEG where we represent the multichannel sleep EEG as a sum of two components: a transient component and an oscillatory component. The transient component is modeled as the sum of a low-frequency signal and a sparse piecewise constant signal. We model the blocks of the multichannel oscillatory component as low-rank arrays. The non-linear signal model proposed in this paper aids in the detection of sleep spindles in multichannel EEG by separating the non-rhythmic transients and the oscillations. Similar signal models have been used to detect sleep spindles directly for single channel [48, 47] and indirectly for multichannel EEG [6].

We estimate the two components of the proposed non-linear signal model by proposing an optimization problem consisting of a convex objective function. We derive a fast matrix-inverse-free algorithm to obtain the solution of the optimization problem. The estimated oscillatory component is then used for detection of sleep spindles. Specifically, an envelope of the bandpass filtered oscillatory component is used to generate a binary signal where 1 indicates a spindle and 0 otherwise. The proposed method does not require the EEG epochs (30 second segments) to be scored for sleep stages in comparison to some state-of-the-art methods [45].

The separation of the transients in the multichannel EEG using the proposed method enables a bandpass filter to display the spindle activity more prominently than simply applying the bandpass filter to the EEG signal. Moreover, separating the transient component effectively attenuates the electrocardiogram (ECG) and other cardiac artifacts from the oscillatory component, which is used for spindle detection. The cardiac artifacts have shown to greatly interfere in spindle density overnight [50]. Since the oscillatory component mostly contains rhythmic oscillations, a low-order bandpass filter suffices for sleep spindle detection [48, 47].

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 formulate a convex objective function for estimating the transient and oscillatory components in the proposed EEG signal model. We derive an iterative algorithm for obtaining the solution to the proposed convex objective function and show how to obtain a binary detection vector using the estimated oscillatory component in Section 3. We show several examples for detection of sleep spindles using the proposed method in Section 4. We evaluate the performance of the proposed detection method using an online EEG database in Section 5 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

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

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

$$\|X\|_* := \mathrm{tr}(X^T X) \tag{4}$$

$$= \sum_{i=1}^{m} \sigma_i(X), \tag{5}$$

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

$$D := \begin{bmatrix} -1 & 1 & & & \\ & -1 & -1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix}. \tag{6}$$

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

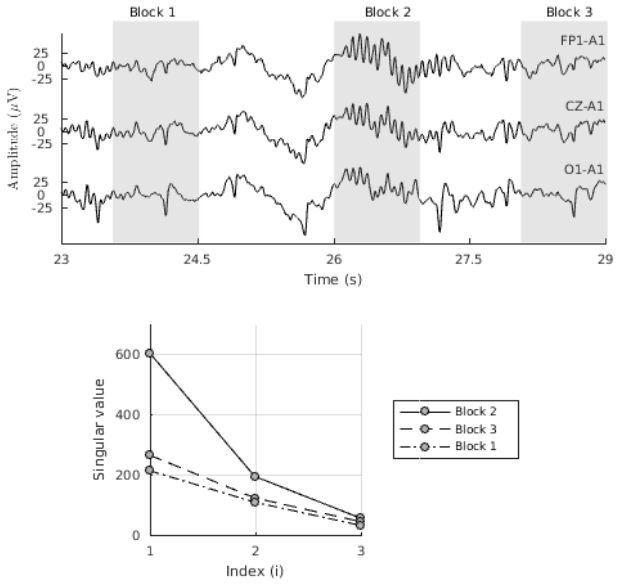

Figure 3: Top: Blocks of multichannel EEG (each of 1 second in length) that contain transients and spindles are highlighted. Bottom: Block 2, which contains expert annotated spindle has higher singular values than Blocks 1 and 3 that do not contain spindles.

[26] for $\lambda > 0$ is defined as

$$\text{soft}(x;\lambda) := \begin{cases} x - \lambda\dfrac{x}{|x|}, & |x| > \lambda \\ 0, & |x| \leqslant \lambda, \end{cases} \qquad x \in \mathbb{R}. \tag{7}$$

Note that the soft-threshold function in (7) 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

$$[T(y)]_n := y^2(n) - y(n-1)\cdot y(n+1). \tag{8}$$

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.

2.2. Block Low-Rank Operator

In order to extract the non-transient oscillations (i.e., sleep spindles) from the multichannel EEG, we propose the following sparse optimization framework. Consider the sample multichannel EEG shown in Fig. 3, with three blocks highlighted (each of 1 second in length). The blocks are highlighted in a non-consecutive fashion to emphasize the difference in their corresponding singular values. Figure 3 shows the singular values corresponding to each of the three blocks. The block that contains the sleep spindles (Block 2) has larger singular values than the blocks which are free of spindle-like waveforms. Consequently, in order to estimate the multichannel oscillatory component we propose a sparse optimization framework wherein we regularize the sum of singular values of the overlapping blocks formed of the input EEG signal. Minimizing the sum of singular values of a matrix results in a low-rank estimate of the matrix [13, 46].

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

$$[\Phi(X)]_i := \begin{pmatrix} x_1(i) & \dots & x_1(i+l-1) \\ & \vdots & \\ x_k(i) & \dots & x_k(i+l-1) \end{pmatrix}, \tag{9}$$

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 $\Phi^T\colon \mathbb{R}^{k\times l\times m} \to \mathbb{R}^{k\times n}$. 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[1].

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

$$\Phi(X) = \Big\{[x_1, x_2], [x_2, x_3]\Big\}. \tag{10}$$

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

$$\mathcal{O}\big(\Phi(X)\big) = [x_1, 2x_2, x_3], \tag{11}$$

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

$$W = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{12}$$

and define

$$\Phi^T(\Phi(X)) = \mathcal{O}\big(\Phi(X)\big)\cdot W \tag{13}$$

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

$$= X. \tag{14}$$

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

$$\mathsf{diag}(W) = \{\underbrace{1,\ldots,1}_{l/2},\underbrace{1/2,\ldots,1/2}_{n-l},\underbrace{1,\ldots,1}_{l/2}\}. \tag{15}$$

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

$$C^* := \arg\min_{C}\left\{\frac{1}{2}\|Y-\Phi^T(C)\|_2^2 + \lambda\sum_{i=1}^{m}\|c_i\|_*,\right\}, \tag{16}$$

where $C = [c_1\ldots,c_m]$, $c_i \in \mathbb{R}^{k\times l}$, $C^* \in \mathbb{R}^{k\times l\times m}$, and $\lambda > 0$ is the regularization parameter. Note that the optimization problem in (16) 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 (16) 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.

The solution to the optimization problem in (16) can be obtained using the iterative shrinkage/thresholding algorithm (ISTA) [7] and its variants. The ISTA algorithm, for the optimization problem in (16), 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 6 singular values respectively.

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:

$$Y := X + S + W, \qquad Y, X, S, W \in \mathbb{R}^{k\times n}, \tag{17}$$

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 [48], 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 [48] and in [47] is similar: the input EEG signal is modeled as a sum of transient and oscillatory components. Note that in this paper we do not estimate the low-frequency transient component of the EEG, which was shown to be of particular interest when detecting K-complexes [48].

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 (17) from the recorded multichannel EEG. To this end, we utilize a sparse optimization framework and propose to solve the following objective function

$$\{X^*, C^*\} = \arg\min_{X,C}\left\{\frac{1}{2}\|Y-X-H(C)\|_2^2 + \lambda_0\sum_{i=1}^{k}\|x_i\|_1 + \lambda_1\sum_{i=1}^{k}\|Dx_i\|_1 + \lambda_2\sum_{i=1}^{m}\|c_i\|_*\right\}, \tag{18}$$

where $X = [x_1,\ldots,x_k]$, $C = [c_1,\ldots,c_m]$, $c_i \in \mathbb{R}^{k\times l}$ and $\lambda_i > 0$ are the regularization parameters. We let $H = \Phi^T$, where Φ is the block low-rank operator (see Sec. 2.2). Recall that D is the first-order difference matrix, as defined in (6), and C is the coefficient array obtained using the operator Φ as defined in (9).

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 [60, 49] and have been shown to model the transient component [48] with relative accuracy.

The nuclear norm on each of the coefficients c_i in (18) 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 (18), we estimate the oscillatory component using the operator $H = \Phi^T$, i.e.,

$$S = \Phi^T(C^*). \tag{19}$$

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

Algorithm 1 McSleep algorithm for solution to (18)

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

$f_1 \leftarrow \frac{1}{\mu} Y + X + D_1$

$f_2 \leftarrow \frac{1}{\mu} H^T Y + C + D_2$

$U \leftarrow f_1 - \frac{1}{\mu+2}\big(f_1 + H f_2\big)$

$V \leftarrow f_2 - \frac{1}{\mu+2} H^T\big(f_1 + H f_2\big)$

$x_i \leftarrow \textbf{soft}\big(\textbf{tvd}(u_i - d_{(1,i)}, \lambda_1/\mu), \lambda_0/\mu\big)$

$[\tilde{U}, \tilde{\Sigma}, \tilde{V}] \leftarrow \textbf{svd}\big(v_i - d_{(2,i)}\big)$

$c_i \leftarrow \tilde{U} \cdot \textbf{soft}(\tilde{\Sigma}, \lambda_2/\mu) \cdot \tilde{V}^T$

$D_1 \leftarrow D_1 - (U - X)$

$D_2 \leftarrow D_2 - (V - C)$

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 (18). 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 (18), 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 (18) under suitable assumptions [29, 10].

We write the objective function (18) using variable-splitting as

$$\arg\min_{X,U,C,V} \Bigg\{ \frac{1}{2}\|Y - U - H(V)\|_2^2 + \lambda_0 \sum_{i=1}^{k} \|x_i\|_1 + \lambda_1 \sum_{i=1}^{k} \|Dx_i\|_1 + \lambda_2 \sum_{i=1}^{m} \|c_i\|_* \Bigg\} \quad \text{such that} \quad U = X, \quad V = C. \tag{20}$$

Using the scaled augmented Lagrangian, minimizing (20) results in solving the following three sub-problems:

$$U, V \leftarrow \arg\min_{U,V} \Bigg\{ \frac{1}{2}\|Y - \big(U + H(V)\big)\|_2^2 + \frac{\mu}{2}\|U - X - D_1\|_2^2 + \frac{\mu}{2}\|V - C - D_2\|_2^2 \Bigg\}, \tag{21a}$$

$$X \leftarrow \arg\min_{X} \Bigg\{ \frac{\mu}{2}\|U - X - D_1\|_2^2 + \lambda_0 \sum_{i=1}^{k} \|x_i\|_1 + \lambda_1 \sum_{i=1}^{k} \|Dx_i\|_1 \Bigg\}, \tag{21b}$$

$$C \leftarrow \arg\min_{C} \Bigg\{ \frac{\mu}{2}\|V - C - D_2\|_2^2 + \lambda_2 \sum_{i=1}^{m} \|c_i\|_* \Bigg\}, \tag{21c}$$

where $\mu > 0$ is the Lagrangian step-size parameter.

The first term in sub-problem (21b) can be written as the energy over each channel of U, X and D_1, i.e.,

$$x_i^* \leftarrow \arg\min_{x_i} \Bigg\{ \sum_{i=1}^{k} \frac{\mu}{2}\|u_i - x_i - d_{(1,i)}\|_2^2 + \lambda_0\|x_i\|_1 + \lambda_1\|Dx_i\|_1 \Bigg\}, \tag{22}$$

with $X^* = [x_1^*, \ldots, x_k^*]$. 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. The solution to (22), for each x_i^*, is readily obtained by applying the fused-lasso method [60] to each channel of the underlying signal, i.e.,

$$x_i^* = \text{soft}(\text{tvd}(u_i - d_{(1,i)}, \lambda_1/\mu), \lambda_0/\mu), \tag{23}$$

where u_i and $d_{(1,i)}$ are the i^{th} channel of U and D_1 respectively. Note that $\text{tvd}(\cdot)$ represents the solution of total variation denoising method [53] obtained using a fast solver [19] and $\text{soft}(\cdot)$ represents the soft-thresholding function (7).

We write the sub-problem (21c) as

$$c_i^* \leftarrow \arg\min_{c_i} \Bigg\{ \sum_{i=1}^{m} \frac{\mu}{2}\|v_i - c_i - d_{(2,i)}\|_2^2 + \lambda_2\|c_i\|_* \Bigg\}, \tag{24}$$

where v_i, c_i and $d_{(2,i)}$ are the i^{th} channel of V, C and D_2 respectively. The solution to (24) is obtained using the singular value thresholding (SVT) algorithm [13], i.e.,

$$[\tilde{U}, \tilde{\Sigma}, \tilde{V}] = \text{svd}(v_i - d_{(2,i)}), \tag{25}$$

$$c_i^* = \tilde{U} \cdot \text{soft}(\tilde{\Sigma}, \lambda_2/\mu) \cdot \tilde{V}^T, \tag{26}$$

where $\text{svd}(\cdot)$ 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 (21a) can be solved exactly using a suitable substitution via the least squares method. Note that the objective function in the sub-problem (21a) is similar to objective function (20a) in

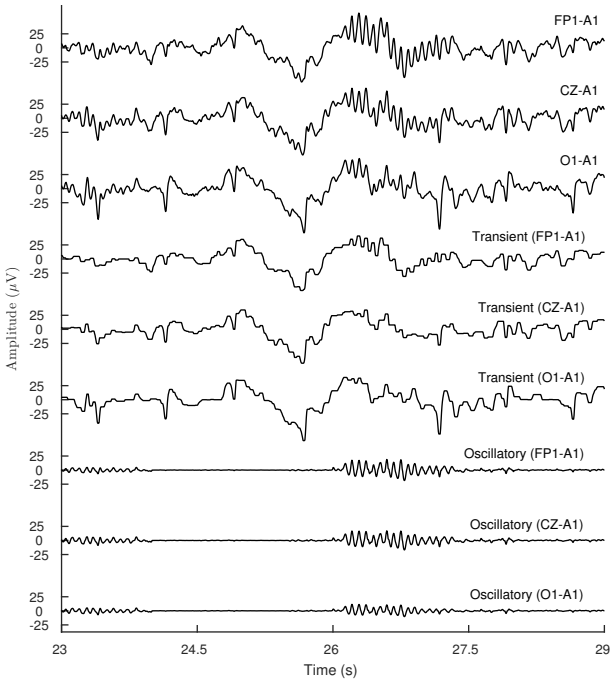

Figure 4: Separation of transients and oscillations using the proposed objective function in (18) for an example EEG segment from DREAMS Database [24]. The transient component is modeled as the sum of a low-frequency signal and a sparse piecewise constant signal.

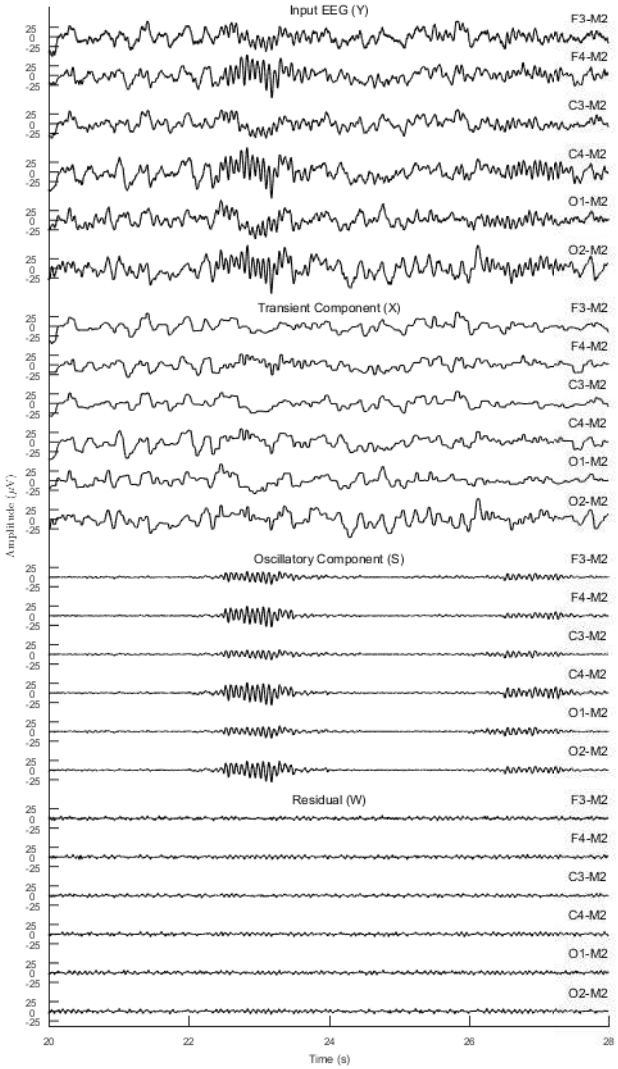

Figure 5: Decomposition of a 6-channel EEG (Y) into its transient (X) and oscillatory (S) components. Also shown is the the residual W, where ($W = Y - X - S$).

[48], and hence a similar derivation can be used for (21a) in this paper. We detail the derivation in Appendix 8.1. The iterative algorithm for (18) is listed in Algorithm 1 and the MATLAB code is made available online[2].

Figure 4 shows the estimated transient and oscillatory components for a three channel EEG (FP1-A1, CZ-A1, O1-A1) from the Devuyst Database[3]. It can be seen in Fig. 4 that the spindles in the three EEG channels are captured by their respective oscillatory components, whereas the non-oscillatory waveforms are captured by the transient components. Also shown in Fig. 5 is the separation of transients and oscillations for a sample 6-channel EEG.

3.4. Detection of Spindles Post Separation of Transients

We use the estimated multi-channel 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 passband of 11 Hz to 16 Hz. 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 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 has been commonly used to obtain an envelop of the bandpass filtered signal for spindle detection [45, 3, 30, 21, 64]. Moreover, a similar framework using a Butterworth bandpass filter and TKEO was used to detect spindles and K-complexes simultaneously [48]. The Teager operator denoted by $T(\cdot)$, is applied to the channel mean of the multi-channel bandpass filtered oscillatory component ($\text{BPF}(S)$) to detect the spindles. Using a constant threshold, we define a binary signal $b_{\text{spindle}}(t)$ as

$$b_{\text{spindle}}(t) := \begin{cases} 1, & T\big(\text{BPF}(S)\big) > c \\ 0, & T\big(\text{BPF}(S)\big) \leqslant c, \end{cases} \tag{27}$$

[2] `https://github.com/aparek/mcsleep.git`

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

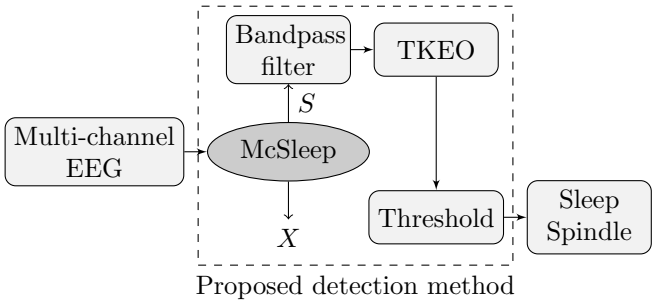

Figure 6: Proposed detection method for multichannel EEG using McSleep as a multichannel transient separation algorithm.

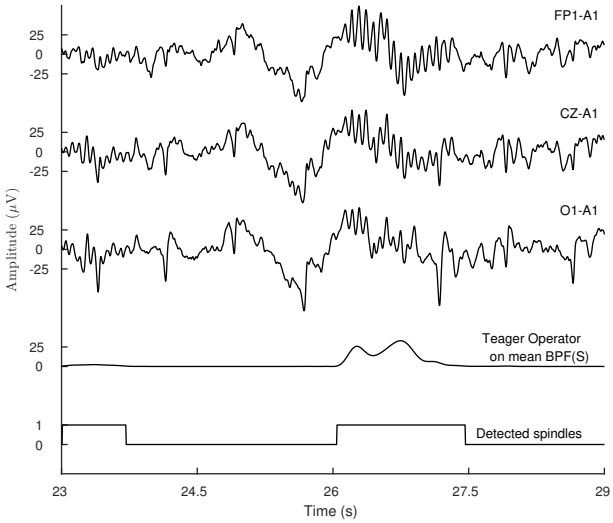

Figure 7: Detection of sleep spindles using the proposed method. The envelope of the bandpass filtered oscillatory component, obtained using the Teager operator, is also shown.

where 1 denotes a spindle present and 0 otherwise.

Figure 6 summarizes the proposed multichannel sleep (McSleep) spindle detection method using the derived transient separation algorithm for sleep EEG. Figure 7 shows the detection of sleep spindles using the proposed McSleep detection method for a 3-channel EEG. Also shown in Fig. 7 is the envelope obtained by applying the TKEO on the mean (across the channels) of the bandpass filtered oscillatory component.

4. Examples

We illustrate the proposed multichannel sleep spindle detection (McSleep) and compare it to other state-of-the-art automated spindle detectors. Recall that the proposed method estimates an oscillatory component that exhibits the block low-rank structure described in Sec. 2.2. The higher the difference in singular values of the blocks that contain spindles and the ones that do not, the better the estimation of the oscillatory component. As such, the proposed method for spindle detection can be run on either the entire EEG montage in a 10-20 system or only the left or the right frontal, central and occipital channels.

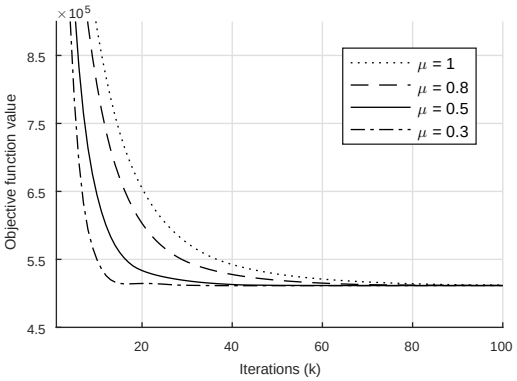

Figure 8: The value of objective function in (18), for each iteration k, is shown for several values of the step-size parameter μ.

For the examples that follow, we run the proposed McSleep method on all the recorded channels of scalp EEG - one central channel (CZ-A1), one frontal channel (FP1-A1) and one occipital channel (O1-A1).

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 (18) 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 the other parameters, enforces the transient component X to be sparse (i.e., with a baseline of zero).

We use $\lambda_0 = 0.3$ and $\lambda_1 = 6.5$ in order to separate the transients from the spindle-like oscillations for an EEG signal at a sampling frequency of 200 Hz. The values for λ_0 and λ_1 were found empirically for the examples that follow. We find that the same λ_0 and λ_1 work well for different EEG signals having the same sampling frequency. Thus for a dataset that contains EEG signals sampled at the same frequency we can preset the values for λ_0 and λ_1. In case an EEG contains relatively more transients, the values of λ_0 and λ_1 may be increased. We set λ_2 to be in the range $(25, 35)$ for an EEG signal sampled at 200 Hz. For EEG signals sampled at a different rate, we scale the parameters $\lambda_0, \lambda_1, \lambda_2$ 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 8 shows the value of the proposed objective function (18) at each iteration for several values of μ. For the examples that follow and for the experiments in Sec. 5, we fix $\mu = 0.5$. Even though a value of $\mu < 0.5$ can be chosen, we observe that for the number of iterations required for convergence (usually 40), all values of $\mu < 0.5$ yield the same objective function value. It can be seen in Fig. 8

that the lowest value of the objective function is achieved after 40 iterations for $\mu = 0.5$ and $\mu = 0.3$.

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 [64, 4, 50]. As a result, we set the block length to be fixed at 1 second with a 50% overlap between consecutive blocks. The threshold c is set to be in the range $(1, 2)$. This range for the threshold value is similar to the range observed in recent studies that use TKEO for spindle detection [44]. For the experimental validation in Sec. 5, the EEG segments that contain body movement artifacts were discarded for analysis visually as in [23]. Furthermore, we discard all detected spindles which are either less than 0.5 seconds or greater than 3 seconds [64].

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 American Academy of Sleep Medicine (AASM) manual [56]. However, increasing number of studies are reporting spindle frequencies to be in a variety of ranges, such as 11–15 Hz [41], 10–17 Hz [64]. 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 spindle detection method can be broken down into two major steps: transient separation and envelope detection using a bandpass filter followed by the Teager operator with a threshold. As a result, studies that may be interested in slow spindles (spindle frequency less than 13 Hz [64]) can set the passband to say 10 Hz to 13 Hz or alternatively set the passband to 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 only λ_2 and c are the 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 datasets in a semi-supervised fashion.

4.2. Example comparison with existing detection methods

We compare the proposed multichannel sleep (McSleep) spindle detection method with the following state-of-the-art automated detectors: Devuyst [23], Wendt [65], Martin [39], and DETOKS [48]. Note that for the proposed McSleep method, we apply the transient separation algorithm on three channels of the scalp EEG (FP1-A1, CZ-A1, O1-A1). We apply the Teager operator on the channel mean of the bandpass filtered oscillatory component to detect spindles.

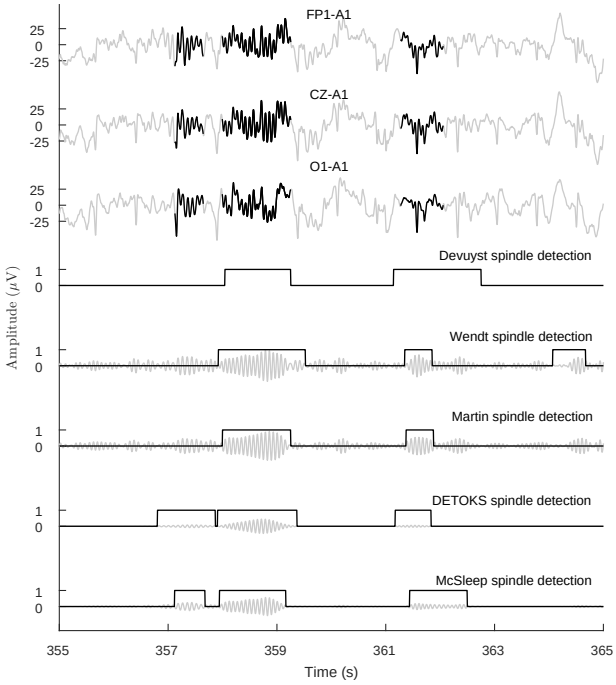

Figure 9: Comparison of proposed McSleep method for spindle detection with existing spindle detectors. Expert annotated sleep spindles are highlighted in black. Bandpass filter output is also shown in background for several methods.

Figure 9 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. 9 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. 9. 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 10 shows the bandpass filtered EEG signal using the filter used by the Wendt algorithm [65]. As seen in Fig. 9, and reported in [48], the Wendt method detects false positive spindles due to the presence of transients. In

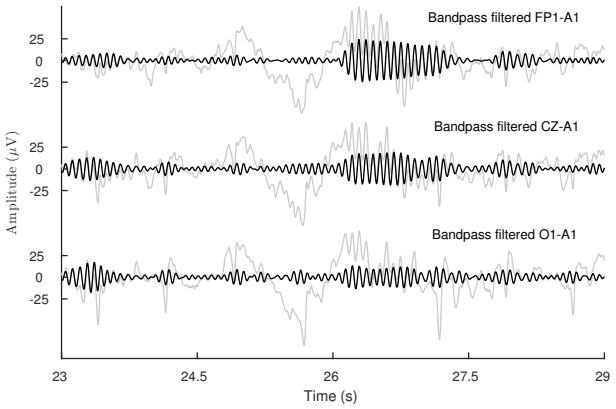

Figure 10: Bandpass filtered sleep EEG using Wendt algorithm [65]. The input EEG is shown in the background. As can be seen, the presence of transients excites the bandpass filter which may lead to false detections.

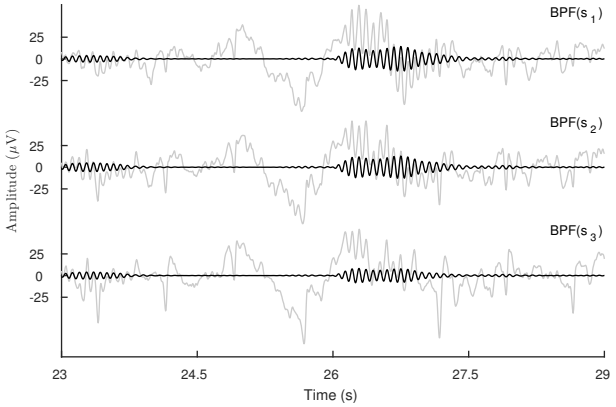

Figure 11: Bandpass filtered oscillatory component estimated using the proposed method. Due to separation of transients, the spindle activity is displayed more prominently.

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. 11.

5. Evaluation of McSleep for Spindle Detection

In this section, we apply the proposed multichannel sleep spindle detection method (McSleep) to an online EEG database wherein the sleep spindles are annotated by experts using only the central (C3-A1) channel. To the best of our knowledge, currently there are no publicly available EEG datasets which contain expert spindle annotation across all channels of the scalp EEG.

5.1. Database and existing automated detectors

In order to evaluate the performance of the proposed method for detecting sleep spindles, we apply the proposed method and the state-of-the-art automated detection algorithms on an online EEG database[4]. We use the following state-of-the-art automated detection algorithms: Wendt [65], Martin [39], Bodizs [8], Wamsley [63], Mölle [33], Devuyst [24], and DETOKS[5] [48]. We refer the reader to [64, 47], and [48] 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.

The online EEG database, used for spindle detection in this paper, 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 evaluated the performance of the proposed method on only 5 of the excerpts as these were scored by both the experts.

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 two methods of analyzing the performance of the automated detectors: by-sample and by-event analysis [64]. Specifically, 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. 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) [40].

In the 'by-event' analysis, a unit of measurement is a single sleep spindle. Specifically, an overlap value is calculated based on the following formula [42, 64]

$$O_{\text{threshold}} := \frac{E \cap D}{E \cup D}, \tag{28}$$

[4] http://www.tcts.fpms.ac.be/~devuyst/#Databases
[5] https://github.com/aparek/detoks

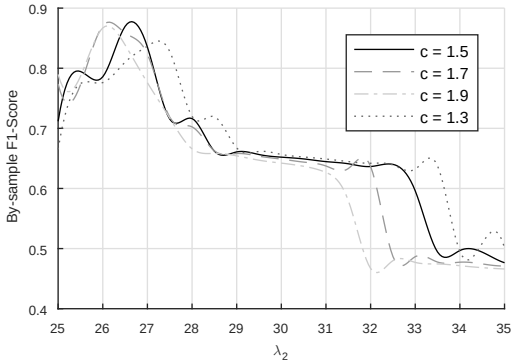

Figure 12: By-sample F_1 score as a function of threshold c and λ_2. For visibility only the c values that yield the highest F_1 scores are shown. The test data is chosen from Excerpt 2 of the online database described in Sec. 5.

where E denotes the expert detection and D denotes the automated detection. A TP is counted if $O_{\text{threshold}} > 0.3$, where 0.3 is user chosen threshold value. All matched event detection (ED) pairs are counted as TP, all un-matched events (E) are counted as FN and all un-matched detections (D) are counted as FP [64]. Since non-spindle events are not meaningful, TN values are not calculated in the by-event analysis. Using the contingency matrix, we calculate the recall and precision values which are future used for calculating the F_1 score. Due to the absence of true negative values, we cannot calculate the MCC score in the by-event analysis. We use an overlap-threshold of 0.3 seconds in order to resolve less-than-perfect overlap and multiple-overlap problems between spindle events and detections. For a pseudo-code of the by-event scoring rule, we refer the reader to [64]. An implementation is provided online with source code of the proposed method.

As a simple example to clarify the differences between the two modes of analyses, consider the following. Let $V = [0, 1, 1, 0, 1, 0]$ represent a binary vector indicating the expert annotation, where 1 represents a spindle and 0 otherwise. Let $A = [0, 1, 1, 0, 0, 1]$ represent the detection via an automated spindle detector. Considering only perfect matches for the by-event analysis, we have TP = 1, FP = 1 and FN = 1, leading to a recall value of 0.5, precision value of 0.5 and an F_1 score of 0.5. In comparison, for the by-sample method of analysis we have TP = 2, FP = 1, and FN = 1, leading to a recall value of 0.67, precision value of 0.67 and an F_1 score of 0.67.

5.3. Parameter Tuning

The evaluation of the proposed McSleep spindle detection method on the Devuyst database requires a set of optimal 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. While the parameters for the McSleep method are described in Sec. 4.1, four key parameters need 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 manually to a fixed value for all the excerpts in the Devuyst database. The fixed value of λ_0 and λ_1 (given in Sec. 4.1) 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 fixed 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 spindle detection method for a grid of values for λ_2 and c. We select the set of parameters that yield the highest F_1 score. The process of selecting λ_2 and c is repeated for another subject in the database. Figure 12 shows the F_1 score as a function of λ_2 and c for a 10 epoch segment for Excerpt 2 from the Devuyst Database. Note that for visual clarity, we only show the F_1 curves associated for four values c. It can be seen from Fig. 12 that the optimal set of parameters is $\lambda_2 \approx 26.5$ and $c = 1.5$. Selecting the optimal parameters based on a small segment of the EEG may seem as over-fitting. As such, we run the proposed method on the entire EEG few times with parameters surrounding the optimal set obtained above.

Table 1: By-event analysis of proposed method for sleep spindle detection as described in Sec. 5. Average values for the F_1 score, recall and precision over 5 excerpts are listed. Standard deviation values are shown in parenthesis.

Methods	By-Event Performance		
	F_1 score	Recall	Precision
Wendt [65]	0.43 (0.11)	0.56 (0.08)	0.36 (0.12)
Martin [39]	0.51 (0.11)	0.51 (0.16)	0.53 (0.05)
Bodizs [8]	0.32 (0.14)	0.70 (0.15)	0.21 (0.14)
Mölle [33]	0.32 (0.26)	0.30 (0.07)	0.34 (0.27)
Devuyst [24]	0.57 (0.07)	0.56 (0.09)	0.58 (0.09)
DETOKS [48]	0.64 (0.03)	0.61 (0.06)	0.69 (0.11)
McSleep	0.67 (0.04)	0.65 (0.05)	0.71 (0.05)

5.4. Results

The average F_1, recall and precision values for the proposed McSleep method in comparison with the other state-of-the-art methods, using the by-event analysis, is listed in Table 1. Table 2 lists the average F_1 score, MCC score, recall and precision values using the by-sample analysis. Detailed statistical measures are listed in Table 3 and Table 4, for the by-event and by-sample methods of analysis respectively, in Appendix 8.2.

The proposed detection method on a 30 minute excerpt from the online 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

Table 2: 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			
	F_1 score	MCC [40]	Recall	Precision
Wendt [65]	0.49 (0.07)	0.47 (0.06)	0.57 (0.08)	0.44 (0.10)
Martin [39]	0.50 (0.08)	0.50 (0.07)	0.43 (0.11)	0.64 (0.06)
Wamsley [63]	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 [48]	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)

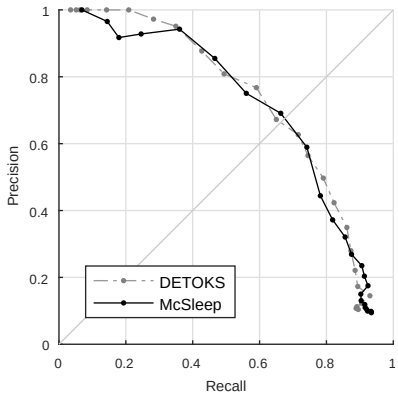

Figure 13: Precision-Recall (PR) curve for proposed McSleep method and the DETOKS method based on the by-event method of analysis. The point closest to the curve from the top-right corner represents the maximum F_1 score obtained.

by bandpass filtering the oscillatory component and applying the TKEO 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 a sample overnight (approx. 8 hours) EEG with 6 channels (as shown in Fig. 5) the proposed method takes on an average 2 minutes. In comparison, a bandpass-filter-based single channel detection method takes on an average 0.5 minutes to 3 minutes, while a transient separation based algorithm, such as DETOKS [48], takes on an average 6 minutes (run in parallel over all the 30 second epochs of all the 6 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 run-times. 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

The proposed McSleep detection method achieved better average F_1 scores compared to other state-of-the-art detectors using by-event analysis. The highest F_1 scores for the by-sample analysis, however, were obtained by the DETOKS [48] method. The Martin [33] and Devuyst [23] methods performed relatively better than other detectors, in terms of the average F_1 scores. The lowest false discovery rates were obtained for the proposed McSleep method. The average false discovery rate (FDR) for the proposed method was 31.3 $\pm$ 0.04% using the by-sample analysis. In comparison, the average FDR for the DETOKS method was 31.8 $\pm$ 0.08% and the average FDR for the Devuyst method was 37.1 $\pm$ 0.05%.

The receiver operating characteristic (ROC) curve is generally used to measure the performance of a binary classifier, such as a spindle detector. However, since the spindle detection problem is asymmetric, i.e., TN values are significantly larger than TP and FP values due to the sparsity of spindle events overnight, the ROC curve is perhaps not the best tool for assessing the performance of a spindle detector [44]. As an alternative, a precision-recall (PR) curve can be used to visually describe the performance of a spindle detector. Since the F_1 scores for the proposed method and the DETOKS method are relatively close for the by-event and the by-sample analysis, we plot the PR curve for these two methods in Fig. 13 to gain more insight into their overall performance. The PR curve is obtained for the by-event analysis by varying the values of the threshold (c) and λ_2 for the proposed method (similar parameters are changed for the DETOKS method). It can be seen in Fig. 13 that the proposed McSleep method and the DETOKS method perform similarly with the McSleep method obtaining a better average F_1 score. A similar PR curve is observed when using the by-sample analysis.

The transient separation algorithm for the proposed McSleep method and the one proposed in DETOKS [48] 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 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 (18) 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 where the gold-standard is expert annotated spindle detection on a single channel of the scalp EEG. However, the proposed method is more precise as it utilizes information from all channels, as seen in Table. 1 and Table 2.

The inter-rater agreement for the dataset we consider

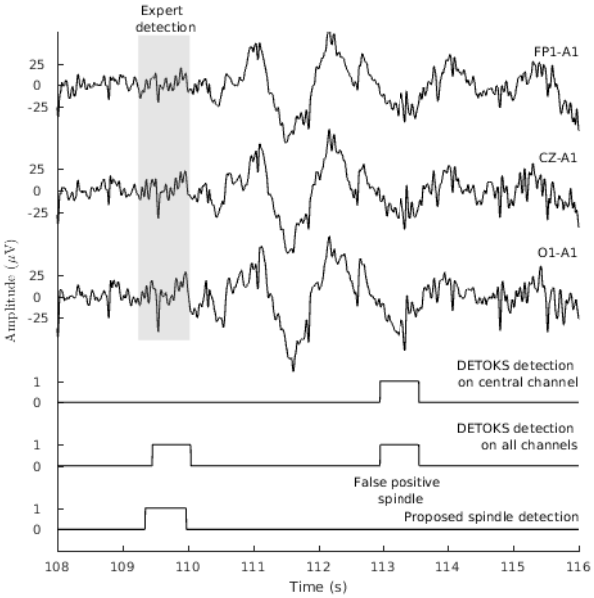

Figure 14: Comparison of the proposed McSleep detection method and the DETOKS detection method. The DETOKS method detects a false positive when run on either a single central channel or on all the channels. The EEG is obtained from excerpt 2 of the online EEG database [24].

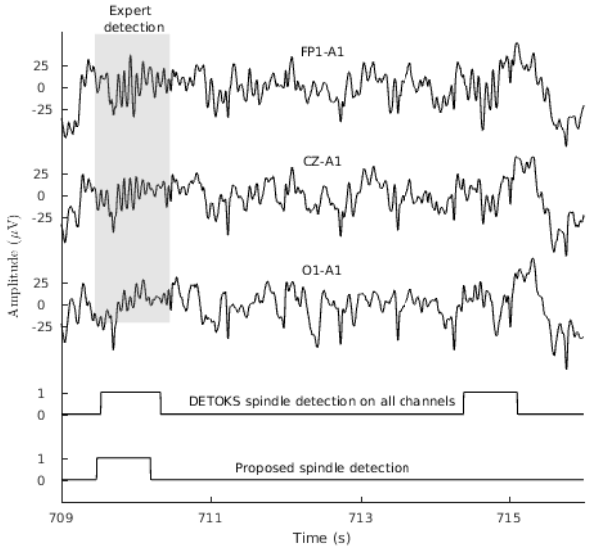

Figure 15: Proposed McSleep detection in comparison with DETOKS method run sequentially on all channels. The proposed detection contains fewer false positive spindles due to better separation of transients.

in this paper is relatively low (on average 44.2% [42]). The relatively low performance of the single channel methods, as well as the proposed method, on the dataset used in this paper may be attributed to the low inter-rater agreement. As a result, using a dataset where the inter-rater agreement is high, may result in a better performance for the automated detectors. We claim that applying the proposed method to a dataset where the experts viewed and annotated all channels of scalp EEG (frontal, central and occipital) will result in a significantly better performance than existing single channel detectors.

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 combine the resulting detections. However, such a multichannel detection method gives rise to two key issues: the high computational complexity associated with running single channel detectors and combining the resulting binary detections. As noted in the preceding subsection, on an average a simple bandpass-filter based spindle detector takes 10 to 100 seconds for a single channel of an overnight EEG, while the MCA based methods take 5 to 6 minutes. The run-time is eight to ten-fold for an overnight EEG acquired using the 10-20 system. On the other hand, the proposed method detects spindles across all the recorded channels of scalp EEG within 2 minutes. Moreover, as spindles may have different amplitudes in different channels, the existing single channel detectors require a separate parameter tuning for each channel. In comparison, with a single set of optimal parameters the proposed method is able to detect spindles across all the channels. Note that the proposed McSleep method can be run on 30 second epochs thus enabling its utility in an online mode[6].

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], [50] and the references therein). Moreover, ensuring that the gold standard adheres to such a voting rule can be challenging. 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 detects a false positive whereas using the union rule the all-channel DETOKS run with separate parameter tuning 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. Moreover, the AASM manual [56] suggests

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

experts to annotate spindles only on the central channel, as such a voting rule wherein higher weight is given to the central channel may not be able to discard the false positive spindle detected in Fig. 14. Similar behavior is observed in another EEG segment shown in Fig. 15.

6. Conclusion

We propose a multichannel transient separation algorithm for detecting sleep spindles (McSleep) simultaneously across the frontal, central and occipital channels of sleep EEG. The proposed method attempts to aid in answering the open question regarding the spatial distribution of sleep spindles. The transient separation algorithms assumes a non-linear signal model for the multichannel EEG. We assume the multichannel EEG to be a sum of a transient and an oscillatory component. The transient component is modeled as a piece-wise constant signal with a baseline of zero whereas the oscillatory component is considered to exhibit block similarity. Specifically, fixed width blocks of the multichannel oscillatory component are of low-rank.

We propose a sparse optimization framework to estimate the transient and oscillatory components. The framework proposed involves minimizing a convex objective function. A fast matrix-inverse-free algorithm is proposed to estimate the transient and oscillatory components. 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.

A fast run-time and F_1 scores in the vicinity of the scores attained by visual experts on a crowd-sourced spindle data [64], 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. Future work includes the setting up of an EEG database wherein sleep spindles are annotated across the scalp in human sleep EEG.

7. 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.*, 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, 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. Konrad, L. Genzel, A. Steiger, M. Dresler and I. Kovács. The hemispheric lateralization of sleep spindles in humans. *Sleep Spindles & Cortical Up States,* (Online), Jan. 2017. doi:10.1556/2059.01.2017.002

[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. Brockmeier and J. Principe. Learning recurrent waveforms within EEGs. *IEEE Trans. Biomed. Eng.*, 63(1):43–54, Jan. 2016. ISSN 1558-2531 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, Jan. 2016. ISSN 0165-0270. doi: 10.1016/j.jneumeth.2015.10.010.

[13] J.-F. Cai, E. J. Candes, and Z. Shen. A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.*, 20 (4):1956–1982, 2010. ISSN 1052-6234. doi:10.1137/080738970.

[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.*, 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. Ste-

nuit, and M. Kerkhofs. Automatic sleep spindle detection in patients with sleep disorders. *Proc. IEEE Int. Conf. Eng. Med. Biol.*, 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.*, 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.*, 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. J. Blinowska. Matching pursuit parametrization of sleep spindles. *Eng. Med. Biol.*, 3:1011–1012, Oct. 1996. doi:10.1109/IEMBS.1996.652685.

[28] P. 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. 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. F. 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.

[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. K. Douglas Leveraging sparsity: A low-rank + sparse decomposition (LR+SD) method for automatic EEG artifact removal *Proc. International Workshop on Sparsity Techniques in Medical Imaging (STMI)*, 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*, 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] 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.

[39] 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.

[40] 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.

[41] S. Noachtar, C. Binnie, J. Ebersole, F. Maguiére, 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

[42] A. Nonclerq, 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, Apr. 2013. ISSN 0165-0270 doi: 10.1016/j.jneumeth.2013.01.015

[43] 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.

[44] 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.

[45] 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.

[46] A. Parekh and I. Selesnick. Enhanced low-rank matrix approximation. *IEEE Signal Process. Lett.*, 23(4):493–497, 2016. ISSN 10709908. doi:10.1109/LSP.2016.2535227.

[47] 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.

[48] 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.

[49] A. Parekh and I. W. Selesnick. Convex fused lasso denoising with non-convex regularization and its use for pulse detection. *Proc. IEEE Signal Process. Med. Biol. Symp.*, Dec. 2015. ISBN 978-1-5090-1350-0 doi:10.1109/SPMB.2015.7405474.

[50] 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.

[51] 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.

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

[53] 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.

[54] 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.

[55] 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.

[56] 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.

[57] 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.

[58] 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.

[59] 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, Jun. 2006. ISSN 0165-0270. doi:10.1016/j.jneumeth.2005.11.006.

[60] 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.

[61] 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.

[62] 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.

[63] 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.

[64] 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.

[65] 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.*, 4250–4253, Aug. 2012. ISSN 1557-170X. doi: 10.1109/EMBC.2012.6346905.

[66] M. E. Wohlleber, A. Parekh, T. Gumb, A. W. Varga, and R. S. Osorio. Reduced spindle frequency is associated with increased cerebrospinal fluid P-Tau in cognitively normal elderly. In *Int. Conf. Sleep Spindl.*, Budapest, 2016.

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

[68] 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.

8. Appendix

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

We derive the solution to the least-squares sub-problem in (21a), which is written below for clarity to the reader.

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

We make the following substitutions

$$\hat{U} = [U, V]^T, \qquad \hat{D} = [D_1, D_2]^T, \tag{29a}$$

$$\hat{X} = [X, C]^T, \qquad M = [I, H], \tag{29b}$$

and re-write the objective function as

$$\hat{U} \leftarrow \arg\min_{\hat{U}} \left\{ \frac{1}{2}\|Y - M\hat{U}\|_2^2 + \frac{\mu}{2}\|\hat{U} - \hat{X} - \hat{D}\|_2^2 \right\}. \tag{30}$$

The solution to (30) can be written explicitly as

$$\hat{U} = \left[M^T M + \mu I\right]^{-1} \left[M^T Y + \mu(\hat{X} + \hat{D})\right]. \tag{31}$$

The inverse in (31) results in inverting a dense matrix consisting of $M^T M$. In order to efficiently compute the explicit solution, we use the Matrix Inverse lemma [67, 62]. As such, the solution in (31) can be written as

$$\left[M^T M + \mu I\right]^{-1} = \frac{1}{\mu}\left(I - M^T\left(\mu I + MM^T\right)^{-1} M\right). \tag{32}$$

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

$$MM^T = [I \quad H] \begin{bmatrix} I \\ H^T \end{bmatrix} \tag{33}$$

$$= 2I. \tag{34}$$

As a result, the inverse in (31) can be written as

$$\left[M^T M + \mu I\right]^{-1} = \frac{1}{\mu}\left(I - \frac{1}{\mu + 2} M^T M\right), \tag{35}$$

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

$$\hat{U} = \frac{1}{\mu} M^T Y + \hat{X} + \hat{D} - \frac{1}{\mu + 2}\left(M^T Y + \mu M^T M(\hat{X} + \hat{D})\right). \tag{36}$$

Combining (36) and (29), we get the following steps for obtaining the solution to the objective function in (21a).

$$f_1 = \frac{1}{\mu} Y + X + D_1 \tag{37a}$$

$$f_2 = \frac{1}{\mu} H^T Y + C + D_2 \tag{37b}$$

$$U = f_1 - \frac{1}{\mu + 2}\left(f_1 + H f_2\right) \tag{37c}$$

$$V = f_2 - \frac{1}{\mu + 2} H^T\left(f_1 + H f_2\right) \tag{37d}$$

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

8.2. Performance evaluation of McSleep for sleep spindle detection

We compare the proposed McSleep spindle detection method with the state-of-the-art automated detectors using the online EEG database [23]. We label the automated detectors as follows: A1 (Wendt et. al [65]), A2 (Martin et. al [39]), A3 (Wamsley et. al [63]), A4 (Bodizs et. al [8]), A5 (Mölle et. al [33]), A6 (Devuyst et. al [23]), A7 (DETOKS [48]) and A8 (proposed McSleep method). The results obtained using the by-event analysis method are summarized in Table 3, while the results obtained using the by-sample analysis are summarized in Table 4 below. For a detailed summary of each of the measure of performance, we refer the reader to [64].

Table 3: By-Event performance of automated sleep spindle detectors in comparison with McSleep. TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative. A8 is the proposed McSleep method.

Measure	Automated Sleep Spindle Detection Algorithms						
	A1 [65]	A2 [39]	A4 [8]	A5 [33]	A6 [23]	A7 [48]	A8
Ex. 1							
TP	77	43	40	52	92	72	78
FP	125	41	70	33	67	21	25
FN	57	91	94	82	42	62	56
Recall	0.57	0.32	0.30	0.39	0.69	0.54	0.58
Precision	0.38	0.51	0.36	0.61	0.58	0.77	0.76
F1	**0.46**	**0.39**	**0.33**	**0.47**	**0.63**	**0.63**	**0.66**
Ex. 2							
TP	45	47	13	41	41	48	52
FP	178	50	44	43	53	24	31
FN	32	30	64	36	36	29	25
Recall	0.58	0.61	0.17	0.53	0.53	0.62	0.68
Precision	0.20	0.48	0.23	0.49	0.44	0.67	0.63
F1	**0.30**	**0.54**	**0.19**	**0.51**	**0.48**	**0.64**	**0.65**
Ex. 3							
TP	18	13	0	0	20	32	26
FP	55	13	0	0	19	33	12
FN	26	31	44	44	24	12	18
Recall	0.41	0.30	0.00	0.00	0.45	0.73	0.59
Precision	0.25	0.50	0.00	0.00	0.51	0.49	0.68
F1	**0.31**	**0.37**	**0.00**	**0.00**	**0.48**	**0.59**	**0.60**
Ex. 5							
TP	58	65	37	0	57	67	80
FP	67	57	63	0	27	19	29
FN	45	38	66	103	46	50	37
Recall	0.56	0.63	0.36	0.00	0.55	0.57	0.68
Precision	0.46	0.53	0.37	0.00	0.68	0.78	0.73
F1	**0.51**	**0.58**	**0.36**	**0.00**	**0.61**	**0.66**	**0.71**
Ex. 6							
TP	76	77	73	69	67	64	72
FP	71	45	130	54	33	23	24
FN	41	40	104	48	50	43	31
Recall	0.65	0.66	0.41	0.59	0.57	0.60	0.70
Precision	0.52	0.63	0.36	0.56	0.67	0.74	0.75
F1	**0.58**	**0.64**	**0.38**	**0.58**	**0.62**	**0.66**	**0.72**

Table 4: By-Sample performance of automated sleep spindle detectors in comparison with the proposed method McSleep. TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative, NPV = Negative Predictive Value. A8 is the proposed McSleep method.

Measure	Automated Sleep Spindle Detection Algorithms							
	A1 [65]	A2 [39]	A3 [63]	A4 [8]	A5 [33]	A6 [23]	A7 [48]	A8
Ex. 1								
TP	8357	4121	2374	9550	4991	9692	9354	8548
TN	156774	164771	166242	137540	164911	159737	162632	16826
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
F_1	**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
Cohen's κ	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
Ex. 2								
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
F_1	**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
Cohen's κ	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
Ex. 3								
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
F_1	**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
Cohen's κ	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
Ex. 5								
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
F_1	**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
Cohen's κ	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
Ex. 6								
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
F_1	**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
Cohen's κ	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