The Generalized Sleep Spindles Detector: A Generative Model Approach on Single–Channel EEGs

Carlos A. Loza[1] and Jose C. Principe[2]

[1] Department of Mathematics, Universidad San Francisco de Quito, Quito, Ecuador
cloza@usfq.edu.ec
[2] Computational NeuroEngineering Laboratory (CNEL), University of Florida, Gainesville, USA
principe@cnel.ufl.edu

Abstract. We propose a data–driven, unsupervised learning framework for one of the hallmarks of stage 2 sleep in the electroencephalogram (EEG)—sleep spindles. Neurophysiological principles and clustering of time series subsequences constitute the underpinnings of methods fully based on a generative latent variable model for single–channel EEG. Learning on the model results in representations that characterize families of sleep spindles. The discriminative embedding transform separates potential micro–events from ongoing background activity. Then, a hierarchical clustering framework exploits Minimum Description Length (MDL) encoding principles to effectively partition the time series into patterns belonging to clusters of different dimensions. The proposed algorithm has only one main hyperparameter due to online model selection and the flexibility provided by cross–correlation operators. Methods are validated on the DREAMS Sleep Spindles database with results that echo previous approaches and clinical findings. Moreover, the learned representations provide a rich parameter space for further applications such as sparse encoding, inference, detection, diagnosis, and modeling.

Keywords: EEG, Generative Model, Representation Learning, Sleep Spindles

1 INTRODUCTION

Sleep spindles constitute the hallmark of stage 2 non-REM sleep. Their generation is attributed to the mutual interaction between GABAergic reticular neurons and excitatory thalamic cells [24] while long–range cortical projections are believed to regulate their temporal synchronization [5]. In terms of behavioral and functional correlates, sleep spindles have been associated to memory consolidation processes [22, 4], cortical development [14], sleep deprivation [7], and are even regarded as potential biomarkers for psychiatric disorders, such as schizophrenia [8, 17]. Therefore, principled detection and modeling are crucial. In the clinical field, EEGers usually utilize scoring rules and norms well documented in the literature [21, 18]; yet, the ever–increasing amount of data and

the advent of machine learning have bolstered the use of automatic sleep spindle detectors as an additional tool for clinicians and neuroscientists [12, 6, 19].

A sleep spindle is defined as a burst in the 11–15 Hz range (sigma band) with duration between 0.5 and 2 s. and a distinctive waxing–wanning envelope. Moreover, as a type of transient events in the EEG, sleep spindles require particular constraints for their detection. Such conditions are derived from both neurophysiology [18] and empirical clinical findings [21]. Automatic detectors explicitly incorporate such constraints into their framework, e.g. amplitude–based thresholds that are cross–validated to ground truth [8, 6]. Classic detectors focus only on timing and amplitude features; yet, as neurophysiological micro–events, sleep spindles are also characterized by duration, frequency, modulation, and, from a generative model stance, by their encoding indexes. All these features can be collectively deemed as representations. A generalized detector should, then, be able to learn such representations in a data–driven manner.

By leveraging the sparse nature of the micro–events, we pose them as samples from a Temporal Marked Point Process (TMPP) that activate elements from a set of vectors (i.e. a dictionary) over time—a generative latent variable model for sleep spindles. A fully unsupervised framework aims to estimate the conditional densities of the latent variables given a set of neuromodulations: generating dictionary (centers of mass in vectors spaces of different dimensionalities), intensity function of the TMPP timings and density of the TMPP marks (amplitudes and indexes). We propose a novel learning algorithm that incorporates neurophysiological constraints and principled techniques for clustering of time series subsequences. We exploit Freeman's theories [9] to restrict the search space of relevant events. Then, a hierarchical clustering algorithm creates a partitioning of the search space by means of Minimum Description Length (MDL) encoding [1]. The result is twofold: learned representations suitable for modeling, and sets of latent variables appropriate for inference.

One of the major advantages of the proposed method is its data–driven nature. Durations of relevant neuromodulations are not limited to a set of user–defined inputs; they are learned on an unsupervised fashion as a result of cross–correlation operators that guarantee flexibility and fine temporal resolution. Also, MDL encoding performs online model selection in a fast, greedy manner. Thus, the proposed algorithm virtually requires only one hyperparameter that is closely related to amplitude–based thresholds of classic sleep spindle detectors.

We validate the methods on the DREAMS Sleep Spindles database [25] and obtain estimates of the representation densities. We compare them with their counterparts from visual scorers and quantify their similarity via the Kullback–Leibler (KL) divergence. We also analyze the effect of the only hyperparameter in terms of receiver operating characteristics (ROC) curves and KL divergences. The results highlight the potential of the proposed method when dealing with principled detection and modeling of sleep spindles. The rest of the paper is organized as follows: Section 2 details the problem to solve while Section 3 presents the methods and rationales behind their choice, Section 4 showcases the results, and, lastly, Section 5 concludes the paper.

2 A problem beyond detection

Let $\tilde{y}[n]$ be a bandpassed single–channel EEG trace that can be decomposed into two time series according to the dynamical regimes of the generating network:

$$\tilde{y}[n] = \begin{cases} y[n] & \text{if Network is Active (Y State)} \\ z[n] & \text{if Network is at Rest (Z State)} \end{cases} \tag{1}$$

where $y[n]$ is the ideal, noiseless component with scale–specific micro–events (sleep spindles), and $z[n]$ is the ongoing, background activity. A mixture model characterizes the probability density of $\tilde{y}[n]$ as:

$$P(\tilde{y}[n]) = p_Y P(\tilde{y}[n]|Y, \Theta_Y) + p_Z P(\tilde{y}[n]|Z, \Theta_Z) \tag{2}$$

where p_Y and p_Z represent the probabilities of states Y and Z parameterized by Θ_Y and Θ_Z, respectively ($p_Y = 1 - p_Z$).

A linear model posits $y[n]$ as the weighted sum of N ideal patterns, $\mathbf{d} \in \mathbf{D}$, shifted over time:

$$y[n] = \sum_{i=1}^{N} \sum_{m=-\infty}^{\infty} \alpha_i \delta[n - \tau_i - m] \mathbf{d}_{\omega_i}[m] \tag{3}$$

where $\delta[n]$ is the Dirac delta function. The elements of the dictionary, $\mathbf{D} = \{\mathbf{d}_{\omega_j}\}_{j=1}^{K}$, do not necessarily have the same dimensionality, i.e. they represent templates with different durations or centers of mass in vector spaces of different dimensions; let such dimensions be the set $\{\phi_i\}_{i=1}^{N}$. Similar generative models with sparsity constraints have been proposed for the auditory nerve [23].

$\tilde{y}[n]$ can be either modeled as the noisy superposition of samples from a TMPP or as the observable variable from a generative latent variable model (Fig. 1) with two distinctive modes: Z, a background component that encodes the spontaneous, disorganized activity of the generating network during rest, and Y, an active component represented as reoccurring transient micro–events that reflect the spatiotemporal synchronization of neuronal assemblies [18, 3]. Freeman posited that Z can be modeled as a Gaussian distribution [10] ($\Theta_Z \triangleq \{\mu_Z, \sigma_Z\}$). Y is the result of joint contributions from latent variables in the form of timings (τ), amplitudes (α), encoding indices (ω), and generating dictionary $\mathbf{D}$, i.e. $\Theta_Y \triangleq \{\tau, \alpha, \omega, \mathbf{D}\}$. The shallow nature of the graph admits the equivalence between latent variables and features or representations. Moreover, $\mathbf{D}$ carries its own features, e.g. duration, frequency, and Q–factor. Consequently, learning on the model can be posed as a type of unsupervised representation learning [2].

Classic sleep spindles detectors usually estimate ϕ and τ leaving $\mathbf{D}$ unaddressed [12, 6, 19]. Given the dictionary $\mathbf{D}$ (where usually $K \geq N$), the goal in analysis, inference, encoding or detection is to estimate the shifts τ, indices ω, and a surrogate of the weights α, for a constrained optimization problem, e.g. Matching Pursuit with the overcomplete Gabor basis as $\mathbf{D}$ [26]. Estimating $\mathbf{D}$ is challenging due to the inherent dynamics of the EEG—unlike classic blind

source separation problems, the relevant sources of $y[n]$ are shift–invariant, non–overlapping and transient. Also, a principled decomposition should discriminate between $y[n]$ and $z[n]$ while preserving the original micro–events. The generalized detector should, then, estimate parameters beyond shifts and durations, i.e. given $\tilde{y}[n]$, the goal in an unsupervised framework is to learn ω, τ, α, ϕ, and $\mathbf{D}$.

3 Methods

The solution to the problem of the previous section is combinatorial in nature. In [15], we propose a solution based on shift–invariant k–means that results in $\phi \in \mathbb{R}^M$, i.e. dictionary elements with predefined duration. Now, we generalize the implementation by, first, restricting the search space and, then, exploiting MDL–based hierarchical clustering to yield prototypical patterns of different durations corresponding to reoccurring sleep spindles in the EEG. Densities of the remaining latent variables naturally arise from the learning process (Fig. 2).

3.1 The Discriminative Embedding Transform

The first step is to isolate $y[n]$ by exploiting the dynamical properties of the EEG. According to Freeman's experimental results, the EEG amplitudes during rest periods resemble a Gaussian distribution, while transitions to work states result in deviations from Gaussianity according to higher–order moments [9]. Both stages, rest and work, alternate in the EEG traces and give rise to transient neuromodulations as a result of the spatiotemporal synchronization of neuronal assemblies [18]. If $\tilde{y}[n]$ is the result of linear filtering, the Gaussian/Non–Gaussian properties during rest/active regimes are preserved for the bandpassed traces.

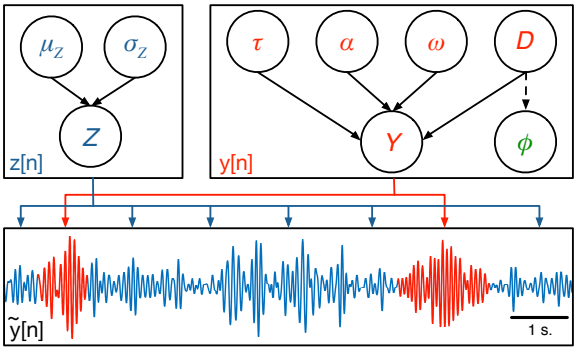


Fig. 1: Generative model of bandpassed single–channel EEG. Z is characterized by the background EEG mean and standard deviation (blue). Y consists of latent variables (red) in the form of timings, weights, indices, and generating dictionary with elements from vector spaces of different dimensions. Durations, ϕ (green), are features from $\mathbf{D}$. Mixing between regimes over time results in $\tilde{y}[n]$.

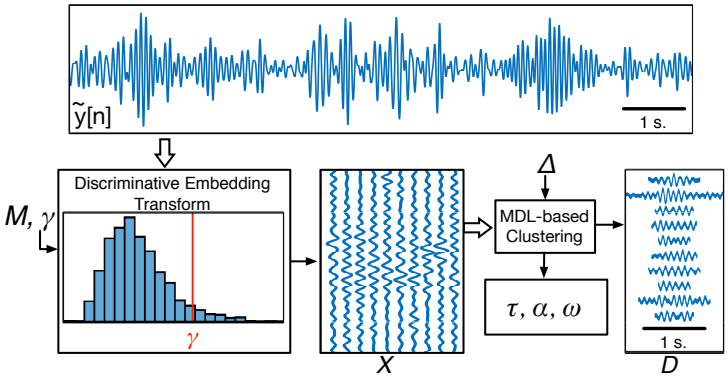


Fig. 2: Proposed generalized sleep spindles detector. Input: Bandpassed single–channel EEG trace, $\tilde{y}[n]$ (top row). Discriminative embedding transform restricts the search space and MDL–based clustering creates a hierarchical structure of patterns from vector spaces of different dimensions. Output: Representations and set of sleep spindles prototypes ($\mathbf{D}$), i.e. clusters, of different durations.

Definition 1. *The M–sample–long subsequence from $\tilde{y}[n]$ centered at the time instance $t = i$ is known as M–snippet:*

$$\tilde{y}_i = \tilde{y}[i - M/2 : i + M/2] \quad \text{s.t} \quad i = M/2, M/2 + 1, \ldots, \eta - M/2 \tag{4}$$

where η is the number of sampled values in $\tilde{y}[n]$. One of the goals of the generalized detector is to discriminate between M–snippets generated by Z (background subsequences) and M–snippets with embedded micro–events generated by Y.

In [16], the embedding transform was introduced as a novel tool to assess stationarity of bandpassed single–channel EEG recordings. In particular, the input time series is non–linearly decomposed into two components based on a surrogate distribution of constrained ℓ_2–norms, β_M:

$$\begin{aligned} \beta_M &= \beta_M(\tilde{y}[n]) \\ &= ||\tilde{y}[\pi_i - M/2 : \pi_i + M/2]||_2 \quad \text{s.t.} \quad \pi_i \in \mathit{\Pi} \end{aligned} \tag{5}$$

What differentiates this approach from a regular embedding is the way the set $\mathit{\Pi}$ is built: the algorithm starts by isolating the indices where relevant modulatory activity is present (peak detection via moving averages or instantaneous amplitudes). Then, the indices corresponding to the remaining unmodulated patterns complete the set $\mathit{\Pi}$. For further details of the algorithm, refer to [16].

After β_M is built, we posit that the M–snippets generated by Z are mapped to a chi–distribution with M degrees of freedom in the β_M space; this density, for large M, results in a Gaussian by the Central Limit Theorem. Conversely, potential relevant sleep spindles, i.e. $y[n]$, are mapped to a second mode in

β_M. M–snippets with corresponding β_M values larger than a hyperparameter threshold γ are collected in the matrix $\mathbf{X} \in \mathbb{R}^{M \times \hat{N}}$:

$$\begin{aligned} \mathbf{x}_i &= \tilde{y}[\pi_i - M/2 : \pi_i + M/2]^T \\ \text{s.t.} \quad &||\tilde{y}[\pi_i - M/2 : \pi_i + M/2]||_2 \geq \gamma \end{aligned} \tag{6}$$

where $\mathbf{x}_i$ is the i-th column of $\mathbf{X}$. In this way, sleep spindles timings are estimated in a similar fashion as classic threshold–based detector strategies [12]:

$$\{\tau_i\}_{i=1}^{\hat{N}} = \{\pi \mid \ ||\tilde{y}_\pi||_2 \geq \gamma \ \text{ and } \ \pi \in \mathit{\Pi}\} \tag{7}$$

In short, $\mathbf{X}$ restricts the search space of $\tilde{y}[n]$ to $\hat{N}$ potential embedded M–sample–long micro–events.

3.2 MDL–based Clustering

After the search space is efficiently restricted, it is necessary to find reoccurring patterns in $\mathbf{X}$. A naive solution would exploit classic clustering algorithms, such as k–means, in the M–dimensional space of the inputs; yet, the solution would include clusters deemed as meaningless [13] due to two main reasons: variable time offsets of patterns from the same cluster, i.e. shift–invariance, and the presence of micro–events of different durations embedded in M–snippets. A plausible solution must address both problems in a principled manner. The former problem is managed via template matching (nearest neighbor search) based on cross-correlations while the latter exploits principles of MDL encoding.

A hierarchical clustering framework greedily selects the number of clusters and estimates reoccurring patterns of different durations by exploiting principles of MDL compression. Let $DL(T)$ be the length of the bit level representation of time series T with length m, i.e. the entropy of T times m.

$$DL(T) = -m \sum_t P(T = t) \log_2 P(T = t) \tag{8}$$

Similarly, the conditional description length of a sequence A after being encoded with a hypothesis H is given by $DL(A|H) = DL(A - H)$, e.g. the cost of the encoding. This principle was applied to hierarchical clustering of time series in [20] under the connotation of *time series epenthesis*. Essentially, the $DL(\cdot)$ operator is a parameter–free tool to evaluate the 3 basic operations in hierarchical clustering: creation of a cluster, assignment of input to existing cluster, and merging of clusters. The cost function BS represents the bits saved after performing the three basic operations, for instance:

BS after creating cluster C from subsequences A and B:

$$BS = DL(A) + DL(B) - DLC(C) \tag{9}$$

where $DLC(C) = DL(H) + \sum_{A \in C} DL(A|H) - \max_{A \in C} DL(A|H)$ is the number of bits needed to represent all subsequences belonging to cluster C and H is the center subsequence of the cluster under consideration.

BS after adding subsequence A to cluster C:

$$BS = DL(A) + DLC(C) - DLC(C') \tag{10}$$

where C' is the new cluster after adding A to C.

BS for merging clusters C_1 and C_2 into new cluster C':

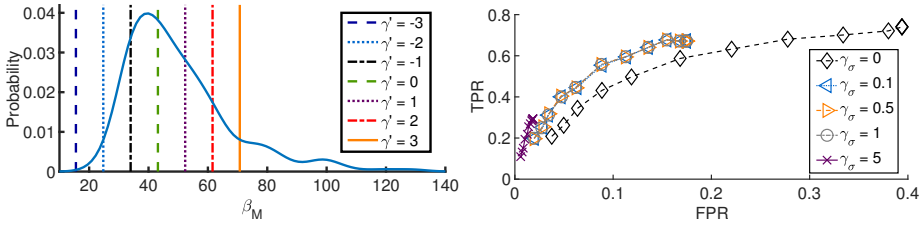
$$BS = DLC(C_1) + DLC(C_2) - DLC(C') \tag{11}$$

The algorithm exposed in [20] utilizes motif discovery algorithms to initialize novel clusters. We propose cross–correlation operations as suitable alternatives for discovering such motifs from $\mathbf{X}$, estimating distances between subsequences and corresponding clusters (α), updating membership vectors (ω), and, ultimately, learning shift–invariant prototypical patterns from vectors spaces of different dimensions. MDL–based encoding basically allows comparison of costs involving clusters of different dimensionalities, which would be prohibitive and unprincipled for the Euclidean distance. The final version of the algorithm evaluates at each step which operation results in the maximal BS and proceeds with such option. Iterations continue until the set $\mathbf{X}$ is exhausted.

The major advantage of this hierarchical clustering framework is twofold: model selection is performed in a greedy manner (K is learned from data) and the resulting sleep spindles prototypes are not restricted to fixed–length patterns. The first hyperparameter, γ, plays the role of an ℓ_2–norm–based threshold of the generalized detector, while the second hyperparameter, $\mathit{\Delta}$, is a set of approximate durations of prospective sleep spindles. Yet, thanks to cross–correlation operators, the final clusters are not necessarily restricted to the elements of $\mathit{\Delta}$; this emphasizes the inherent adaptive nature of the proposed framework. Moreover, $\mathit{\Delta}$ can be chosen in a principled manner according to the rhythm under consideration [18]. Therefore, the proposed generalized detector is a virtually one–hyperparameter learning mechanism that sequentially selects in a greedy manner the operation that maximizes the bits saved among the possible δ–dimensional inputs, where $\delta \in \mathit{\Delta}$ and $\delta \leq M$.

4 Results

Sleep spindles latent variables are estimated on the DREAMS Sleep Spindles database [25] for the available 8 subjects. Single–channel (either CZ-A1 or C3-A1), 30–minute–long EEG traces were made available with their corresponding visual scorings of sleep spindles (sampling frequencies ranging from 50 to 200 Hz). The sigma band is isolated using Butterworth filters with quality factor (ratio of central frequency and bandwidth) $Q \approx 2$. M is set equal to the sample equivalent of 1.5 seconds. $\mathit{\Delta}$ is set to [0.5:0.1:1.5] sec. according to scoring criteria of sleep spindles [21, 18, 19]. For proper validation and estimation of density similarities, we only utilize the scores from one (visual scorer 1) of the two experts due to the strong bias in durations from visual scorer 2—zero standard deviations and mean durations of 1 s.



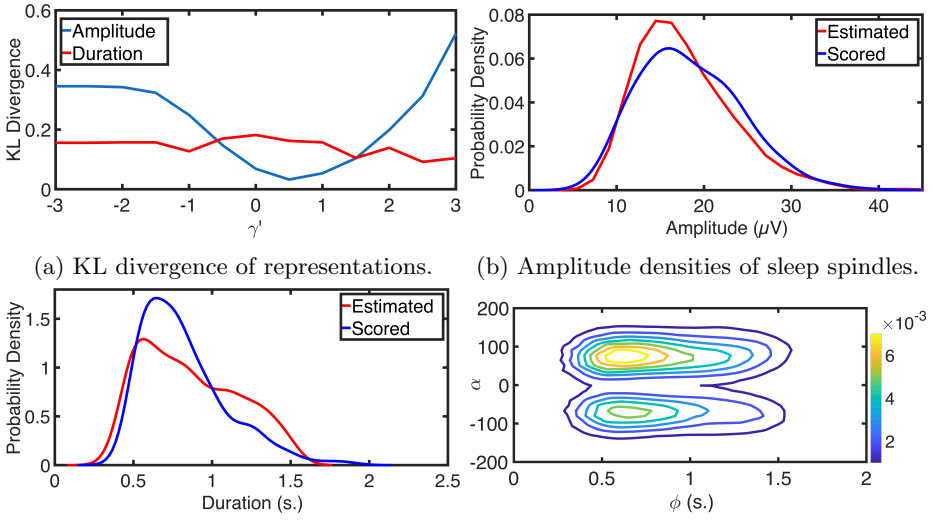
(a) Embedding Transform (Subject 2). Vertical lines indicate estimated γ thresholds according to $\gamma = \mu_{Z_M} + \gamma' \times \sigma_{Z_M}$.

(b) ROC curves over γ' for several values of γ_σ. For any ROC curve, points correspond to γ' values in the interval [-3:0.5:3].

Fig. 3: Effects of the Embedding Transform–based threshold, γ, over the M–snippet domain and performance in terms of sleep spindles detection.

ROC curves are a good starting point to validate estimated timings and durations in terms of True Positive Rates (TPR) and False Positive Rates (FPR). Fig. 3 shows ROC curves (grand average) over $\gamma = \mu_{Z_M} + \gamma' \times \sigma_{Z_M}$ where μ_{Z_M} and σ_{Z_M} are the mean and standard deviation of the set of M–snippets generated from Z, respectively. We test γ' in the range [-3:0.5:3]. Also to reduce the FPR, the sigma index [11, 12] is exploited to reject alpha intrusions and EMG interference. In particular, the sigma index threshold, γ_σ, is a lower bound for the ratio between powers in the sigma band and neighboring rhythms. Fig. 3b compares ROC curves for several values of γ_σ while Fig. 3a indicates the estimated γ thresholds in β_M for a sample subject. In general, the ROC curves are very robust for a wide range of γ_σ and quickly saturate for $2 \leq \gamma' \leq 3$. Best cases correspond to a global sensitivity of 67.7% and FPR = 0.154 compared to 70.2% and 0.264 from the original report [6], respectively.

Next, we validate the estimated amplitudes and durations. Here it is worth noting that classic sleep spindles detectors (and visual scorers) do not share the generative nature of the generalized detector and, hence, define amplitude as the absolute value of the peak amplitude during a micro–event. Duration, as previously noted, is not a latent variable, but rather a representation inferred from the learned dictionary; hence, durations parameterize—and act as surrogate features of—$\mathbf{D}$. Similarity between scored and estimated representations is assessed via the KL divergence; in this way, we generalize the concepts classic detectors gauge with TPRs and FPRs. Amplitude representations display a local minimum KL divergence at $\gamma' = 0.5$, while durations are relatively unaffected (Fig. 4a). This implies that the generalized detector robustly learns the duration density, and therefore, is able to robustly learn the generating dictionary (at least in terms of surrogate features). Conversely, KL divergences of the amplitudes provide a novel criterion for threshold selection based on representation densities from a generative model instead of classic performance measures from ROC curves. Learned densities (Fig. 4b and 4c) echo the experimental results of Purcell et al. [19] in a massive study of sleep spindles characterization.



(a) KL divergence of representations. (b) Amplitude densities of sleep spindles.

(c) Duration densities of sleep spindles. (d) Joint density of weights and duration.

Fig. 4: Learned densities from single–channel EEGs and associated measures. (a) KL divergence between learned and scored representations. (b) and (c) Estimated densities for sleep spindles amplitude and duration ($\gamma' = 0.5$). (d) Contour plot of bivariate joint density of weights and duration from generative model ($\gamma' = 0.5$). Color bar represents probability density—a surrogate of the membership index latent variable, ω.

The remaining latent variables and representations are analyzed next. The weights, α, represent the distance between detected micro–events and their corresponding clusters or prototypes, whereas $\mathbf{D}$ can be partially characterized by the duration ϕ. Their joint bivariate density clearly indicates two main modes corresponding to sleep spindles in the range [0.5, 1] seconds and symmetric marginal weight densities (Fig. 4d). Membership indices, ω, shape the probability density over the weight–duration space and, hence, define the bimodal density. The latent variable model can easily sample from this distribution to generate micro–events, and hence, simulate single–channel EEG traces with embedded sleep spindles.

Table 1 summarizes some statistics from the learned latent variables for each subject. σ_Z is the estimated standard deviation of the background component (a measure of the rest RMS of this rhythm), the median of the inter micro–event interval (IMEI) characterizes the shifts τ in a similar manner as interspike intervals for units. Averages of weight magnitudes are also reported (α densities were bimodal and symmetric around zero). Lastly, median durations and number of clusters, K, parameterize the learned dictionaries succinctly. The measures of Table 1 can be further exploited for inference, e.g. sleep disorder diagnosis; large–scale modeling; and encoding—Matching Pursuit–based detectors [26] with an ensemble dictionary as suitable alternative to wavelets or Gabor bases.

Table 1: Estimated parameters from learned representations. SX denotes Subject X according to DREAMS Sleep Spindles database. $\mu_Z = 0$ for bandpassed traces.

	Subject							
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**
σ_Z (μV)	4.08	5.07	3.45	4.32	3.29	3.51	4.42	3.49
med(IMEI) (s.)	4.44	5.10	4.94	8.09	4.16	4.89	4.07	3.30
avg($\lvert\alpha\rvert$)	73.35	104.04	36.45	100.82	73.49	82.91	87.07	73.09
med(ϕ) (s.)	0.87	1.00	1.06	0.71	0.66	0.69	0.77	0.80
K	62	53	55	36	49	42	48	61

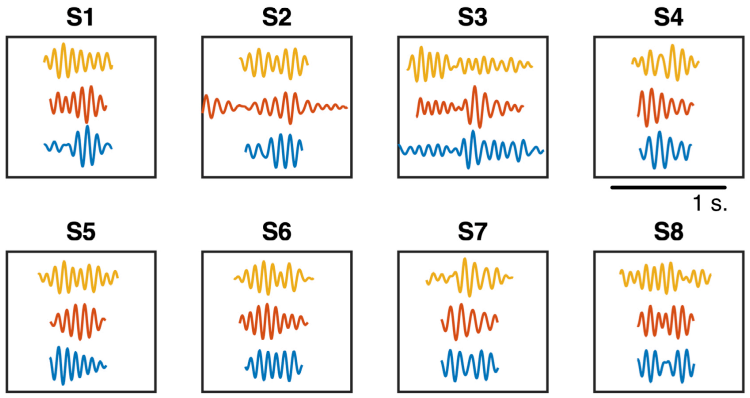


Fig. 5: Sample prototypical sleep spindles learned from single–channel EEGs. Clusters with highest number of micro–events assignments are shown.($\gamma' = 0.5$).

Fig. 5 depicts some of the learned clusters for each subject. The proposed one–hyperparameter method discovers in a data–driven manner prototypical sleep spindles with a wide range of temporal supports and modulatory patterns. Moreover, $\mathbf{D}$ can be characterized by its own features or representations beyond duration, e.g. frequency, number of oscillations, symmetry, and Q–factor. This opens the door to rich parameter spaces where inference and modeling are appealing. For instance, Fig. 6 characterizes the dictionaries via their power spectral density (PSD). The smoothing effect in the cluster estimation resembles ensemble averages in spectral estimation, which helps mitigate the bias.

5 Conclusion

A generative latent variable model for sleep spindles generalizes classic detectors to be able to learn representations from two physiological regimes in an unsupervised manner. The proposed methodology discovers bases from vector spaces of different dimensions, i.e. clusters of time series subsequences of different durations. The one–hyperparameter algorithm efficiently learns features that can be further exploited by clinicians as tools in encoding, detection, inference, and

modeling. Future work includes iterative estimation of the latent variables in a Bayesian framework where the EM algorithm results advantageous.

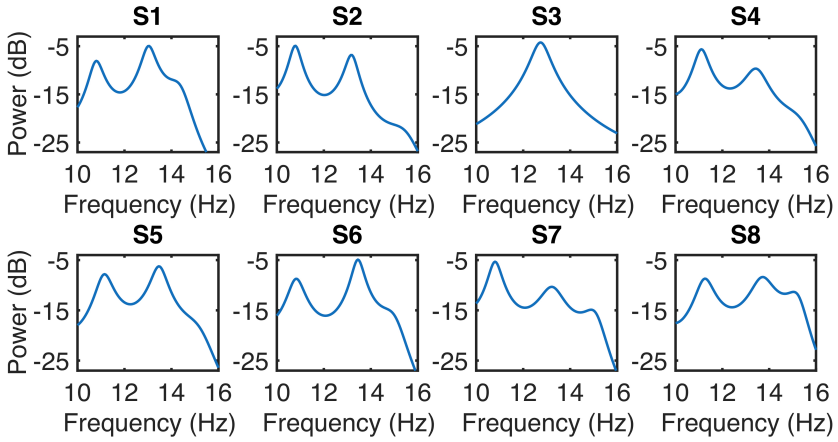


Fig. 6: Estimated Power Spectral Densities of sleep spindles prototypes (Autoregressive model–based estimation with AIC model selection, $\gamma' = 0.5$).

References

1. Barron, A., Rissanen, J., Yu, B.: The minimum description length principle in coding and modeling. IEEE Transactions on Information Theory 44(6), 2743–2760 (1998)
2. Bengio, Y., Courville, A., Vincent, P.: Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence 35(8), 1798–1828 (2013)
3. Buzsáki, G., Anastassiou, C.A., Koch, C.: The origin of extracellular fields and currentsEEG, ECoG, LFP and spikes. Nature reviews neuroscience 13(6), 407 (2012)
4. Clemens, Z., Fabo, D., Halasz, P.: Overnight verbal memory retention correlates with the number of sleep spindles. Neuroscience 132(2), 529–535 (2005)
5. Contreras, D., Destexhe, A., Sejnowski, T.J., Steriade, M.: Control of spatiotemporal coherence of a thalamic oscillation by corticothalamic feedback. Science 274(5288), 771–774 (1996)
6. Devuyst, S., Dutoit, T., Stenuit, P., Kerkhofs, M.: Automatic sleep spindles detectionoverview and development of a standard proposal assessment method. In: Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE. pp. 1713–1716. IEEE (2011)
7. Dijk, D.J., Hayes, B., Czeisler, C.A.: Dynamics of electroencephalographic sleep spindles and slow wave activity in men: effect of sleep deprivation. Brain research 626(1-2), 190–199 (1993)
8. Ferrarelli, F., Huber, R., Peterson, M.J., Massimini, M., Murphy, M., Riedner, B.A., Watson, A., Bria, P., Tononi, G.: Reduced sleep spindle activity in schizophrenia patients. American Journal of Psychiatry 164(3), 483–492 (2007)

9. Freeman, W., Quiroga, R.Q.: Imaging brain function with EEG: advanced temporal and spatial analysis of electroencephalographic signals. Springer Science & Business Media (2012)
10. Freeman, W.J.: Mass action in the nervous system (1975)
11. Huupponen, E., Värri, A., Himanen, S.L., Hasan, J., Lehtokangas, M., Saarinen, J.: Optimization of sigma amplitude threshold in sleep spindle detection. Journal of sleep research 9(4), 327–334 (2000)
12. Huupponen, E., Gómez-Herrero, G., Saastamoinen, A., Värri, A., Hasan, J., Himanen, S.L.: Development and comparison of four sleep spindle detection methods. Artificial intelligence in medicine 40(3), 157–170 (2007)
13. Keogh, E., Lin, J.: Clustering of time-series subsequences is meaningless: implications for previous and future research. Knowledge and information systems 8(2), 154–177 (2005)
14. Khazipov, R., Sirota, A., Leinekugel, X., Holmes, G.L., Ben-Ari, Y., Buzsáki, G.: Early motor activity drives spindle bursts in the developing somatosensory cortex. Nature 432(7018), 758 (2004)
15. Loza, C.A., Okun, M.S., Príncipe, J.C.: A marked point process framework for extracellular electrical potentials. Frontiers in systems neuroscience 11, 95 (2017)
16. Loza, C.A., Principe, J.C.: The embedding transform. a novel analysis of non-stationarity in the EEG. In: Engineering in Medicine and Biology Society (EMBC), 2018 IEEE 40th Annual International Conference of the. p. to appear. IEEE (2018)
17. Manoach, D.S., Pan, J.Q., Purcell, S.M., Stickgold, R.: Reduced sleep spindles in schizophrenia: a treatable endophenotype that links risk genes to impaired cognition? Biological psychiatry 80(8), 599–608 (2016)
18. Niedermeyer, E., da Silva, F.L.: Electroencephalography: basic principles, clinical applications, and related fields. Lippincott Williams & Wilkins (2005)
19. Purcell, S., Manoach, D., Demanuele, C., Cade, B., Mariani, S., Cox, R., Panagiotaropoulou, G., Saxena, R., Pan, J., Smoller, J., et al.: Characterizing sleep spindles in 11,630 individuals from the national sleep research resource. Nature communications 8, 15930 (2017)
20. Rakthanmanon, T., Keogh, E.J., Lonardi, S., Evans, S.: Time series epenthesis: Clustering time series streams requires ignoring some data. In: Data Mining (ICDM), 2011 IEEE 11th International Conference on. pp. 547–556. IEEE (2011)
21. Rechtschaffen, A., Kales, A., University of California, L.A.B.I.S., Network, N.N.I.: A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. Publication, Brain Information Service/Brain Research Institute, University of California (1968)
22. Schabus, M., Gruber, G., Parapatics, S., Sauter, C., Klösch, G., Anderer, P., Klimesch, W., Saletu, B., Zeitlhofer, J.: Sleep spindles and their significance for declarative memory consolidation. Sleep 27(8), 1479–1485 (2004)
23. Smith, E.C., Lewicki, M.S.: Learning efficient auditory codes using spikes predicts cochlear filters. In: Advances in Neural Information Processing Systems. pp. 1289–1296 (2005)
24. Steriade, M., McCormick, D.A., Sejnowski, T.J.: Thalamocortical oscillations in the sleeping and aroused brain. Science 262(5134), 679–685 (1993)
25. TCTS Lab: The DREAMS sleep spindles database (2011), `http://www.tcts.fpms.ac.be/~devuyst/Databases/DatabaseSpindles/`
26. Żygierewicz, J., Blinowska, K.J., Durka, P.J., Szelenberger, W., Niemcewicz, S., Androsiuk, W.: High resolution study of sleep spindles. Clinical Neurophysiology 110(12), 2136–2147 (1999)