International Journal of Neural Systems, Vol. 0, No. 0 (2019) 1–??
© World Scientific Publishing Company

Computationally-efficient algorithm for real-time absence seizure detection in wearable electroencephalography

Jonathan Dan
STADIUS - ESAT KU Leuven, Leuven, Belgium
Byteflies, Antwerp, Belgium

Benjamin Vandendriessche
ECSE - Case Western Reserve University, Cleveland, Ohio, United States of America
Byteflies, Antwerp, Belgium

Wim Van Paesschen
Neurology - UZ Leuven, Leuven, Belgium
Department of Neurology - KU Leuven, Leuven, Belgium

Dorien Weckhuysen
Neurology - Kempenhaeghe, Heeze, Netherlands

Alexander Bertrand
STADIUS - ESAT KU Leuven, Leuven, Belgium
E-mail: alexander.bertrand@kuleuven.be

Advances in electroencephalography (EEG) equipment now allow monitoring of people with epilepsy in their daily-life environment. The large volumes of data that can be collected from long-term out-of-clinic monitoring require novel algorithms to process the recordings on board of the device to identify and log or transmit only relevant data epochs. Existing seizure-detection algorithms are generally designed for post-processing purposes, so that memory and computing power are rarely considered as constraints. We propose a novel multi-channel EEG signal processing method for automated absence seizure detection which is specifically designed to run on a microcontroller with minimal memory and processing power. It is based on a linear multi-channel filter that is precomputed offline in a data-driven fashion based on the spatial-temporal signature of the seizure and peak interference statistics. At run-time, the algorithm requires only standard linear filtering operations, which are cheap and efficient to compute, in particular on microcontrollers with a multiply-accumulate unit. For validation, a dataset of eight patients with juvenile absence epilepsy was collected. Patients were equipped with a 20-channel mobile EEG unit and discharged for a day-long recording. The algorithm achieves a median of 0.5 false detections per day at 95% sensitivity. We compare our algorithm with state-of-the-art absence seizure detection algorithms and conclude it performs on par with these at a much lower computational cost.

Keywords: epilepsy; electroencephalography; absence seizures; signal processing; automated seizure detection; wearable EEG; embedded systems

1. Introduction

Epilepsy is one of the most common severely disabling brain conditions, affecting over 46 million people worldwide.[1] The clinical utility of monitoring epileptic seizures has been amply documented.[2] For example, it is used for optimizing antiepileptic therapy and for the development of new drugs. Until recently, long-term monitoring of epilepsy relied mostly

on patient-reported outcomes and occasional short duration electroencephalographic (EEG) recording in hospital-based epilepsy monitoring units to document epileptic discharges. Such in-hospital recordings rarely contain actual electroclinical seizures as these are relatively unlikely to occur during the very short recording sessions.[3] For a small proportion of people with epilepsy, typically those selected for surgical treatment, this can eventually be complemented by additional (one to several day-long) video-EEG recordings in specially equipped hospital units.[2] Epilepsy monitoring units confine recordings within a hospital room, imposing restriction on patients movement and activity. These hospital recordings also incur significant expenses.[4]

Epilepsy is classified in different syndromes and seizure types. Typical absence seizures are generalized seizures that occur in people with juvenile absence epilepsy and childhood absence epilepsy syndromes. These seizures are characterized by sudden impairment of consciousness lasting a few seconds.[5] The seizures are associated with a bilateral, synchronous, and symmetrical EEG discharge of 3Hz spike and wave complexes.[6] Given the lack of distinctive movement or other clinical manifestations, it is difficult for people with absence seizures or their caretakers to accurately detect and count seizures, and report them to the clinician. An ambulatory EEG study in people with absence seizures found that only 6% of seizures lasting more than three seconds were reported.[7] Due to this severe underreporting, a wearable EEG device that automatically detects and logs absence seizures in daily life situations would be a valuable asset for the people with epilepsy and the clinician.

The emergence of miniature EEG devices that can record EEG outside a hospital or lab environment enables ambulatory measurements in a real-life setting.[8–12] This allows long-term recordings, which have a much higher chance of capturing seizures.[13] In addition, seizures resulting from epileptic triggers that would not occur in the hospital can be recorded, sleep patterns are not disturbed, and the neurologist has access to a more representative recording of the patient's seizures. In order for these new EEG devices to be accepted and worn by people with epilepsy outside the hospital for prolonged periods of time, the devices need to be designed to be comfortable, miniature, wearable and discreet. Such EEG devices could be used to monitor people with epilepsy over long periods of time, generating large amounts of data. In turn, this would require automated data analysis to extract relevant biomarkers from large datasets as the amount of manual annotation work would not be manageable realistically by an epileptologist. Furthermore, the recorded EEG data have to be processed on the device itself in order to minimize the amount of data that have to be logged within the device or transmitted to the cloud. This would ensure a sufficiently long battery lifetime.

Many seizure detection algorithms have been proposed for use in the clinic.[14,15] Most of these are not built for a specific seizure type and are operated offline on data collected during routine clinical observation in epilepsy monitoring units. Only a few algorithms have been designed specifically for people with absence seizures.[16–20] In even fewer of these studies, algorithms were designed to run on wearable systems. In contrast to algorithms developed to run in a hospital without strict computational constraints (as they can rely on powerful servers connected to the hospital network),[15] algorithms for wearable systems must meet with strict storage, computing memory, and computing power constraints. The challenge of developing energy efficient seizure detection algorithms has also highlighted by teams working on long-term intracranial EEG for use in implanted devices.[21–23]

In this paper we propose a fully automated seizure detection algorithm specifically designed to run on a microcontroller. The algorithm is based on a linear filter that is designed in a data-driven fashion to maximally amplify absence seizure signals while optimally attenuating peak interference. This framework is inspired by work from Wouters et al. in spike sorting where a template matching algorithm is used.[24] Wouters et al. propose to optimize a data-driven filter in terms of separation of target signal and peak interferences, thereby maximizing the signal-to-peak interference ratio. This is an adaptation of classical data-driven filters that optimize signal-to-noise (SNR) ratio. It is then applied as a linear filter-and-sum operation with finite impulse response (FIR) filters. FIR filtering is a standard signal processing operation that can be efficiently implemented in low-power hardware, in particular when a dedicated multiply-accumulate (MAC) unit is available.

The paper is organized as follows. The data collection and the data-driven filter design are presented in section 2. The results of the proposed algorithm are reported and compared to other state-of-the-art absence seizure detection algorithms in section 3. The results are discussed in section 4 and conclusions are drawn in section 5.

2. Materials & methods

2.1. *Data collection and annotation*

The main dataset used in this study consisted of recordings performed in out-of-clinic environments. An additional validation dataset was assembled from retrospective recordings obtained in a hospital setting to validate the generalizability of the proposed algorithm to new independent data sets with additional patients and different recording equipment.

For the main dataset, EEG signals were obtained from eight patients with refractory juvenile absence epilepsy recruited at the epilepsy reference center of the UZ Leuven university hospital (Belgium) (Table 1). Inclusion criteria were age of 18 or above, and diagnosis of refractory juvenile absence epilepsy ascertained by an expert neurologist. Exclusion criteria for data analysis were the absence of recorded absence seizures. Participants were equipped at the hospital with a 20 channel Medatec BrainWalker3 (Braine-le-Chteau, Belgium) portable EEG amplifier (10-20 system) sampling at 200Hz and then discharged for ambulatory continuous EEG recording until the next day. During the recording, they were allowed to proceed with their daily life activities (the only restriction was not washing their hair).

For the independent validation dataset, EEG recordings from 17 patients with refractory juvenile absence epilepsy who visited the epilepsy monitoring unit during the period 2016-2019 were aggregated (Table 2). The patients were recorded within the video-EEG unit for a day-long monitoring using a 20 channel BrainRT Brainbox 1042 EEG recorder (Kontich, Belgium) sampling at 250Hz. Throughout those recordings, the patients stayed in the video-EEG monitoring room. The first eight patients from table 2 are the same as as the patients in table 1.

Absence seizures, defined as electroencephalographic generalized spike and wave discharges , were annotated by an expert epileptologist from UZ Leuven. It is noted that epileptiform graphoelements with a duration of less than three seconds are generally associated with interictal activity.[25] It depends on the targeted use case whether such shorter events should be flagged by a detection algorithm or not, and so whether their detection should be treated as true or false positives. In this work, this choice was avoided by excluding all epileptiform segments with a duration shorter than three seconds from the analysis.

The study was approved by the local ethics committee and written informed consent was obtained from all participants.

2.2. *Seizure detection algorithm*

Our proposed seizure detection algorithm is a pre-trained spatio-temporal filter, implemented as a filter-and-sum pipeline; followed by thresholding on the power of the single-channel output signal. This approach is illustrated in Figure 1.

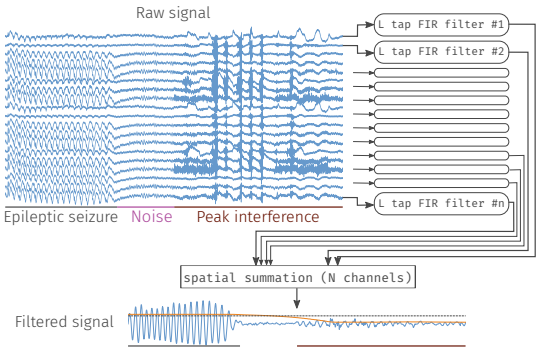


Figure 1. Illustration of the seizure detection algorithm. Bandpass filtered multichannel EEG signal (top left, blue) is fed to a pre-trained (see Fig. 2) spatio-temporal filter (right) implemented as a filter-and-sum pipeline where each FIR filter has L filter coefficients ('taps'). This results in a single-channel output (bottom, blue). The time-varying standard deviation (RMS value) of this filter output signal is computed (bottom, orange). All samples above a set threshold (bottom, black) are labeled as seizures.

The training of the filter coefficients is done in a two-step approach. First, peak interferences are detected by training and applying a purely spatial filter that maximizes the seizure power to non-seizure power ratio. The 40 minutes of highest power

Table 1. Patients with juvenile absence epilepsy recruited for 24h out-of-clinic EEG monitoring. Sex: F (female), M (male); Age in years; Seizure types: A (absence seizure), GTCS (generalized tonic–clonic seizure), M (myoclonic seizure); Antiepileptic medication: CLB (clobazam), CZP (clonazepam), LCM(lacosamide), LEV (levetiracetam), LTG (lamotrigine), PER (perampanel), TPM (topiramate), VPA (valproic acid); Median seizure duration in seconds

Patient	Sex	Age	Seizure types	Antiepileptic medication	Recorded seizures	Seizure duration	Recording duration
1	F	49	A, GTCS, M	CZP, LEV, TPM, VPA	0	NA	05:53:30
2	F	25	A, GTCS	LEV, LTG, PER	41	26	22:43:47
3	M	18	A, GTCS	LCM, VPA	8	10	21:45:00
4	M	24	A	LTG, PER, VPA	0	NA	NA
5	F	42	A, GTCS	CLZ, LCM, VPA	5	8	21:18:20
6	F	22	A, GTCS	LTG	5	4	21:55:33
7	F	20	A, GTCS	CLB, LCM	24	13	21:43:40
8	F	24	A, GTCS, M	CZP, LTG, VPA	18	9	21:50:40

Table 2. Patients with absence epilepsy recruited for 24h in-hospital EEG monitoring. Sex: F (female), M (male); Age in years; Seizure types: A (absence seizure), GTCS (generalized tonic-clonic seizure), M (myoclonic seizure); Antiepileptic medication: CLB (clobazam), CZP (clonazepam), ESM (ethosuximide), LCM(lacosamide), LEV (levetiracetam), LTG (lamotrigine), PER (perampanel), TPM (topiramate), VPA (valproic acid)

Patient	Sex	Age	Seizure types	Antiepileptic medication	Recorded seizures	Recording duration
1	F	49	A, GTCS, M	CZP, LEV, TPM, VPA	5	21:59:25
2	F	25	A, GTCS	LEV, LTG, PER	0	20:42:50
3	M	18	A, GTCS	LCM, VPA	0	20:16:50
4	M	24	A	LTG, PER, VPA	5	20:38:04
5	F	43	A, GTCS	CLZ, LCM, VPA	0	14:01:40
6	F	22	A, GTCS	LTG	0	20:56:13
7	F	20	GTCS	CLB, LCM	0	19:02:21
8	F	24	A, GTCS, M	CZP, LTG, VPA	0	20:12:47
9	M	28	A, GTCS, M	ESM, VPA	20	19:54:40
10	F	30	A, GTCS	ESM, LEV, LTG	2	21:48:35
11	F	41	A, GTCS	ESM, LCM, LEV	10	19:44:12
12	M	24	A	VPA	0	20:34:56
13	F	32	A	LTG	0	21:22:48
14	F	47	A, GTCS	LEV, TPM, VPA	88	21:16:33
15	F	49	A, GTCS	TPM, VPA	6	19:40:05
16	F	48	A	LTG, TPM	40	19:54:04
17	F	33	A, GTCS	LEV, LTG	43	21:54:56

of the single-channel output (excluding seizure segments) are labeled as peak interference. Then, a second spatial-temporal data-driven filter that maximizes the seizure to peak interference ratio (SPIR) is trained. This filter can be adapted in terms of number of channels and number of time lags. A detailed explanation of the algorithm is given in the following subsections.

2.2.1. *Data preprocessing*

The data were high-pass and low-pass filtered with a fourth-order Butterworth filter with a cut-off frequency of 0.5Hz and 25Hz, respectively. The data were then downsampled to 50Hz and re-referenced to a longitudinal bipolar montage[26] resulting in 18 channels. Segments with a root mean

square (RMS) amplitude of more than 400μV over a window of 100ms were considered as non-electroencephalographic (i.e. not of brain origin) and excluded from the analysis along with 1.5 seconds of data before and after each excluded segment of 100ms.

2.2.2. *Data-driven filter design*

The EEG signal in channel k is modeled as

$$y_k(t) = s_k(t) + n_k(t) + i_k(t)$$

where $s_k(t)$ corresponds to seizure-related EEG activity, $n_k(t)$ corresponds to background EEG or noise components that are not related to the seizure [a], and $i_k(t)$ corresponds to peak interference, i.e. artifacts that occur sparsely in time with a high peak amplitude, e.g. due to chewing or eye blinking. Peak interferers are typically the main source of false positives in a threshold-based seizure detection algorithm, whereas the noise floor denoted by $n_k(t)$ generally does not trigger a detection. Note that $s_k(t)$ and $i_k(t)$ are sparse processes, i.e. the signals are zero most of the time, whereas the noise floor signal $n_k(t)$ is continuously active. The raw data in figure 1 show periods where seizure activity, noise, and peak interference (chewing artifact) are respectively dominant.

Let $\mathbf{y}(t) \in \mathbb{R}^N$ denote a N-dimensional vector containing the sample at time t collected at N EEG channels, i.e. $\mathbf{y}(t) = [y_1(t) \ \dots \ y_N]^T$, where T denotes the transpose operator. The vectors $\mathbf{s}(t)$, $\mathbf{n}(t)$, and $\mathbf{i}(t)$ are defined similarly, such that $\mathbf{y}(t) = \mathbf{s}(t) + \mathbf{n}(t) + \mathbf{i}(t)$. The N EEG channels are then linearly combined into a single-channel output signal $o(t)$.

$$o(t) = \mathbf{w}^T\mathbf{y}(t) \tag{1}$$

where $\mathbf{w} \in \mathbb{R}^N$ contains the combination weights. From a signal processing point of view, $\mathbf{w}$ acts as a spatial filter as it linearly combines different EEG channels at different positions on the scalp. Our goal is to find the optimal set of filter coefficients such that the filter output signal $o(t)$ has a maximal amplitude if a seizure is present, while suppressing the noise floor and peak interferers as much as possible. In other words, the filter $\mathbf{w}$ is optimized in a data-driven fashion to maximize the SPIR of $o(t)$ over a training set, i.e. solving

$$\max_{\mathbf{w}} \frac{E_s\{(\mathbf{w}^T\mathbf{s}(t))^2\}}{E_i\{(\mathbf{w}^T(\mathbf{i}(t)+\mathbf{n}(t)))^2\}} \tag{2}$$

where $E_s\{.\}$ denotes the expected value operator evaluated over epochs in which a seizure is present, and $E_i\{.\}$ denotes the expected value operator evaluated over signal segments during which a peak-interferer is active.

Equation 2 is equal to:

$$\max_{\mathbf{w}} \frac{\mathbf{w}^T\mathbf{R_s}\mathbf{w}}{\mathbf{w}^T\mathbf{R_{i+n}}\mathbf{w}} \tag{3}$$

where $\mathbf{R_s} = E_s\{\mathbf{s}(t)\mathbf{s}(t)^T\}$ is the seizure covariance matrix and $\mathbf{R_{i+n}} = E_i\{(\mathbf{i}(t)+\mathbf{n}(t))(\mathbf{i}(t)+\mathbf{n}(t))^T\}$, is the peak-interference-plus-noise covariance matrix. In subsection 2.2.3, we will explain how these covariance matrices are estimated from the data. It can be shown[27] that the solution of the maximization problem defined in equation 3 is the eigenvector corresponding to the largest eigenvalue of the matrix $\mathbf{R_{i+n}^{-1}}\mathbf{R_s}$.

The filter described above is a purely spatial filter. It can be expanded to a causal spatio-temporal filter by creating a buffer of L samples for each channel and stacking all buffered (time-lagged) samples in a single vector $\tilde{\mathbf{y}}(t) = col\{\tilde{\mathbf{y}}_\mathbf{1}(t), \dots, \tilde{\mathbf{y}}_\mathbf{N}(t)\}$ where $\tilde{\mathbf{y}}_\mathbf{k}(t) = [y_k(t), y_k(t-1), \dots, y_k(t-L+1)]^T$ and $col\{.\}$ denotes a columnwise stacking. The output signal $o(t)$ is given by equation 1 where $\mathbf{w}$ is replaced by $\tilde{\mathbf{w}} \in \mathbb{R}^{LN}$ and $\mathbf{y}(t)$ is replaced by $\tilde{\mathbf{y}}(t) \in \mathbb{R}^{LN}$. This then corresponds to a filter-and-sum operation as depicted in Fig. 1, where each channel is filtered with a channel-specific L-taps FIR filter, followed by a summation across channels. The filtering operation that produces one output sample of $o(t)$ is obtained by LN multiplications and LN additions.

In this spatio-temporal extension, the covariance matrices in equation 3 are replaced with their spatio-temporal generalizations, i.e., $\mathbf{R_s} = E_s\{\tilde{\mathbf{s}}(t)\tilde{\mathbf{s}}(t)^T\}$ and $\mathbf{R_{i+n}} = E_i\{(\tilde{\mathbf{i}}(t)+\tilde{\mathbf{n}}(t))(\tilde{\mathbf{i}}(t)+\tilde{\mathbf{n}}(t))^T\}$. In the remainder of the paper, we shall always assume the spatio-temporal extension unless otherwise specified, and omit the $\tilde{}$ for notational convenience.

[a] In the remaining of the paper, we will refer to $n_k(t)$ as the 'noise' component. The term noise should be read in the signal processing terminology as non-target signal components, i.e., signal dynamics that are not related to a seizure.

2.2.3. *Covariance matrix estimation*

In order to compute the optimal filter $\mathbf{w}$, we need to estimate the two covariance matrices $\mathbf{R_s}$ and $\mathbf{R_{i+n}}$ on a training data set, where the training data can be either patient-specific or patient-independent (see subsection 2.3).

In our implementation of the algorithm, the seizure covariance matrix $\mathbf{R_s}$ was computed as [b]

$$\mathbf{R_s} \approx \frac{1}{|\mathcal{S}|} \sum_{t \in \mathcal{S}} \mathbf{y}(t)\mathbf{y}(t)^T \tag{4}$$

where $\mathcal{S}$ is the set of all samples that are part of epochs that were marked as "seizure" by an expert epileptologist. Similarly, the peak-interference covariance matrix $\mathbf{R_{i+n}}$ was computed as

$$\mathbf{R_{i+n}} \approx \frac{1}{|\mathcal{I}|} \sum_{t \in \mathcal{I}} \mathbf{y}(t)\mathbf{y}(t)^T$$

where $\mathcal{I}$ is the set of all samples that are part of epochs that contain a peak interference. As opposed to $\mathcal{S}$, the set of samples that belong to $\mathcal{I}$ are identified in an automatic fashion. To this end, again, a data-driven filtering technique was used.

The signal was modeled as a linear combination of a seizure time series $\mathbf{s}(t)$ and a seizure-free time series $\mathbf{f}(t)$ (including noise-only and peak-interference-plus-noise segments). A spatial filter $\mathbf{p}$ was then computed as the eigenvector belonging to the largest eigenvalue of $\mathbf{R_f^{-1}R_s}$, where the seizure covariance matrix $\mathbf{R_s}$ was computed as in equation 4. The seizure-free covariance matrix $\mathbf{R_f}$ was computed on the rest of the data. Note that in this prior step, we do not use any temporal filtering, i.e. L is set to 1 resulting in a purely spatial filter. This filter was then used to produce a single-channel output given by $\mathbf{p}^T\mathbf{y}(t)$. The RMS over three seconds of the output was computed. The 40 minutes of highest RMS during seizure-free epochs were labeled as peak interference which then form the set $\mathcal{I}$. The set is formed by ranking the signal by highest power, selecting the sample with the highest power, extending the selection to three seconds around the sample and repeating until a total of 40 minutes are selected. This captures most of the peak interferers that appear during a 24-hour recording such as, e.g. chewing artifacts, eye blinking, speech artifacts, head motion, etc while being short enough such that high power interferers dominate the set $\mathcal{I}$. The process leading to the identification of the peak interferers and the training of the max-SPIR filter is shown in figure 2.

2.2.4. *Regularization*

The inclusion of time lags substantially increases the dimension of the covariance matrices, possibly making them ill-conditioned due to the redundancy in the entries of $\mathbf{y}(t)$. For this reason a regularization scheme is required. Wouters et al.[28] proposed an effective regularization scheme for a template-matching-filter that optimizes a similar cost function as in equation 3. The regularization is obtained by projecting the data on a subspace containing the main principal components of the denominator covariance matrix along with the template itself (to represent the target signal). In this work we adapted the method to our problem formulation. The data were projected on a subspace defined as the span of the principal components of the peak interference $\mathbf{R_{i+n}}$ and seizure $\mathbf{R_s}$ covariance matrices. The principal components with the largest eigenvalues of $\mathbf{R_{i+n}}$ and accounting for 90% of the variance in the interference segments were retained. The principal components with the largest eigenvalues of $\mathbf{R_s}$ accounting for 95% of the variance in the seizure segments were also retained. These two sets of principal component vectors were then combined and orthogonalized by placing them in the columns of a new matrix $\mathbf{M}$ on which a singular value decomposition is applied to find an orthogonal basis for the subspace. Let $\mathbf{U} = [\mathbf{u}_1 \dots \mathbf{u}_K]$ denote the matrix, the columns of which consist of the K left singular vectors corresponding to the largest singular values of $\mathbf{M}$, where the cut-off K is chosen such that the cumulated sum of these singular values is at least 99% of the sum of all singular values. The matrix $\mathbf{U}$ is then used as a compression matrix on the data, i.e.,

$$\mathbf{y_c}(t) = \mathbf{U}^T\mathbf{y}(t)$$

where $\mathbf{y_c}(t) \in \mathbb{R}^K$. The optimal compressed filter $\mathbf{w_c}$ is then given as the eigenvector of $(\mathbf{R_{i+n,c}^{-1}})\mathbf{R_{s,c}}$,

[b] Note that, since the noise floor is always present, the resulting covariance matrix is actually an estimate for $E_s(\mathbf{s}(t) + \mathbf{n}(t))(\mathbf{s}(t) + \mathbf{n}(t))^T = \mathbf{R_s} + \mathbf{R_n}$. If necessary, $\mathbf{R_n}$ can be estimated over noise-only epochs (without seizures or peak interferers) and subtracted from the estimate of $\mathbf{R_s}$. However, since absence seizures have a much higher amplitude than the noise floor generated by background EEG, we have not applied this correction in our implementation.

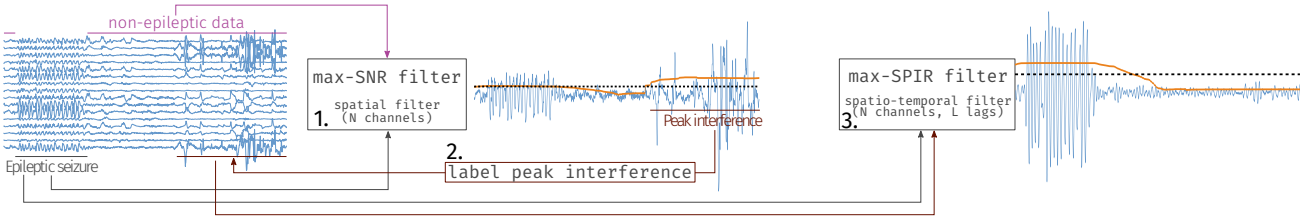


Figure 2. Illustration of the three-stage process that is used to train the spatio-temporal max-SPIR filter. (1) Bandpass filtered multichannel EEG (left, blue) is used to train a purely spatial max-SNR filter. The filter optimally amplifies the epochs of epileptic activity and attenuates those of non-epileptic data. (2) Examples of peak interference are automatically labeled based on the filtered data with the highest RMS power during non-epileptic epochs. (3) These peak interferences and the epileptic activity epochs are used to train a max-SPIR spatio-temporal filter. In contrast to the max-SNR filter, the max-SPIR filter will focus on suppressing peak interferers (which cause the majority of the threshold crossings) more so than suppressing sub-threshold noise.

where $\mathbf{R_{i+n,c}} = \mathbf{U}^T\mathbf{R_{i+n}}\mathbf{U}$ and $\mathbf{R_{s,c}} = \mathbf{U}^T\mathbf{R_s}\mathbf{U}$ correspond to the compressed matrices. The filter output is then defined as $o(t) = \mathbf{w_c}^T\mathbf{y_c}(t)$ which is equivalent to an uncompressed filtering $o(t) = \mathbf{w}^T\mathbf{y}(t)$ with $\mathbf{w} = \mathbf{U}\mathbf{w_c}$. It is noted that this compression is only applied with the purpose of regularization during training to obtain a better (uncompressed) filter $\mathbf{w}$. During operation of the algorithm (at test time), we always apply the full filter $\mathbf{w}$ on the uncompressed data, as the compression with $\mathbf{U}$ requires more computations than the filtering with $\mathbf{w}$.

2.2.5. *Seizure detection*

The power of the single-channel output $o(t)$ was used to detect samples corresponding to a seizure. The RMS of $o(t)$ over three seconds was calculated as this duration is commonly used to define an absence seizure.[5] A threshold was then applied to the running RMS signal, selected depending on the desired sensitivity (see Section 3). This binary classification (above or below threshold) classified samples as seizure or non-seizure.

Several rules were applied to the output of the binary classification to assess the performance of the classification. Samples labeled as seizures and occurring less than 1.5 seconds before the start or after the end of a seizure were not counted as false positive. This was done to account for the settling time of the rolling RMS. False positives occurring less than 30 seconds apart were merged as a single false positives. These rules were applied both to our proposed algorithm and to the algorithms used as benchmark.

2.3. *Cross-validation*

When training and testing the seizure detection algorithm in a patient-specific paradigm, care was taken to separate the dataset in independent training and testing sets. The seizure epochs were split in two folds of equal size (referred to as seizure folds). The rest of the data were split in four folds (referred to as non-seizure folds) of continuous data (each non-seizure fold contained approximately six hours of data). Each training fold used one fold of seizure epochs and three folds of non-seizure data. The obtained filter was then evaluated on the remaining seizure fold and remaining non-seizure fold. This process was repeated for all possible combinations between seizure and non-seizure training folds. An illustration of the cross-validation scheme for the patient-specific case is given in figure 3.

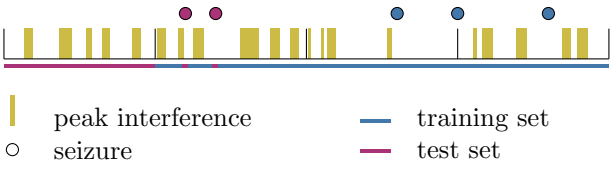


Figure 3. Patient-specific cross-validation. The seizures are split in two folds. The remaining data are split in four folds. One seizure fold and three non-seizure folds are used for training. The remaining seizure fold and remaining non-seizure fold are used for testing. This process is repeated until all combinations of training and testing folds are covered.

In the patient-independent paradigm, a leave-one-patient-out approach was used for cross-

validation. The seizure $\mathbf{R_s}$ and peak interference $\mathbf{R_{i+n}}$ covariance matrices were calculated for each patient, normalized with respect to their trace, and then averaged across patients to obtain the pair ($\mathbf{R_s}$, $\mathbf{R_{i+n}}$) that is used for the filter design in equation 3. The process was repeated by systematically leaving one patient out of the training of the filter and testing on that patient.

As a final validation, the patient-independent filter trained on the patients discharged for out-of-clinic monitoring was evaluated on a second dataset with additional patients and a different EEG device, recorded in hospital. This was used as a validation test to show that the algorithm is generalizable to independent EEG recordings.

2.4. *Channel selection*

When evaluating the performance of the algorithm, we investigated the effect of the number of channels (N) on the performance. To this effect a channel selection procedure that selects N out of all available channels was used. It is based on a greedy forward selection. In the first iteration, 18 single-channel SPIR filters are trained and evaluated. The channel with the best performance (in terms of false positives for a sensitivity of 95%) was selected ($N = 1$). In the next iteration, all subsets of two channels that include the previously selected channel were trained and evaluated ($N = 2$). The process was repeated until N channels were selected. Note that the evaluation of the best channels was done on the training fold rather than the test fold. Although this may introduce an overfitting bias to the training fold, the reported performance (on the test fold) is fair in the sense that the channel selection is not optimized with respect to the test fold. The use of a separate validation fold for channel selection was not possible due to the small amount of seizures in some patients.

2.5. *Evaluation metrics*

A varying seizure detection threshold was used to map sensitivity, i.e. the percentage of seizures that are detected by the algorithm, in function of false positive count. This threshold was applied to the running mean of the RMS of the single-channel output $o(t)$ as explained in subsection 2.2.5. A measure of false positive rate was used in place of specificity because seizures are very rare events and specificity of the algorithm does not correctly assess the clinical utility of seizure detection algorithms.[29]

In the patient-specific paradigm, the false positives were averaged over the eight different test folds, to calculate an average false positive rate per day.

Because energy efficiency of the algorithm is important when operating in real time on board of a wearable device, the algorithm was also evaluated in terms of memory usage and computing complexity. The computing complexity was calculated as the number of operations (summations and multiplications required by the algorithm). Single operations that do not depend on the number of channels or the number of samples were not counted.

2.6. *Comparison with the state of the art*

We compared our method with two state-of-the-art algorithms for absence seizure detection using wearable EEG.[16,17] Kjaer et al.[16] proposed a support-vector machine (SVM) based classifier for use with a single-channel EEG setup. The algorithm extracts many features based both on amplitude and frequency before applying a non-linear SVM classifier. The algorithm suggested by Xanthopoulos et al.[17] is based on a multichannel wavelet-based algorithm followed by a thresholding operation. These algorithms were re-implemented as truthfully as possible based on the original publications and evaluated on our dataset. A more detailed description of the similarities and differences between these algorithms and our proposed method is given in the discussion.

3. Results

All results presented in this section are for the main (ambulatory) dataset, unless mentioned otherwise. Data from all but two subjects could be used for analysis. Data from patient 4 were lost due to a recording error. Patient 1 had no seizures during the recording time. These two subjects with no recorded seizures were excluded from the analysis.

3.1. *Performance analysis*

Figure 4 shows the false positive rate in function of seizure detection sensitivity when all ($N = 18$) bipolar channels were used and the number of time lags was set to $L = 25$ in a patient-specific paradigm. A

distribution of the false positive rate across the subjects is shown for each sensitivity level. The median number of false positives per day for a sensitivity of 95% is 0.5. Note that we report the mean number of false positives across the two seizure folds. This is why the number of false positives per day are reported as multiples of 0.5, despite the fact that we never record longer than a day. In this context, a value of 0.5 false positives per day within a patient means that the filter trained on seizure fold 1 generates 0 false positives per day, while the filter trained on seizure fold 2 generates 1 false positive per day. When displaying results in a boxplot the following conventions are used. The median is represented as a bright square. The box extends from the first to the third quartile. The whiskers extend to 1.5 times the interquartile range $(1.5 * Q3 - Q1)$. All points outside this range are represented as large dots.

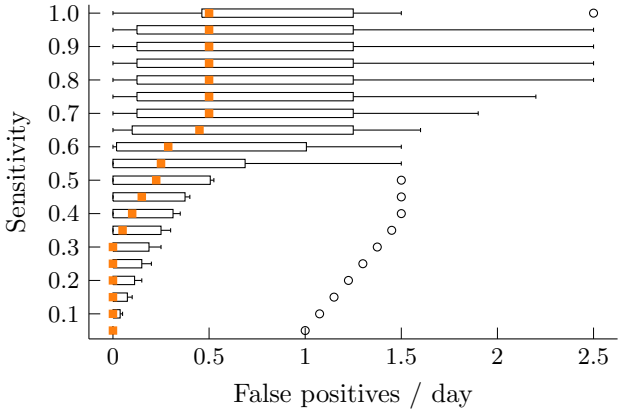


Figure 4. Boxplot (over the different subjects) of seizure detection sensitivity as a function of false positives per day when applying the data-driven filter with all ($N = 18$) channels and $L = 25$ time lags in a patient-specific paradigm. Median is represented by an orange square, outliers by a black circle.

Table 3 shows the number of false positives when detecting all seizures and when detecting all but one seizure in the different patients when all ($N = 18$) bipolar channels were used and the number of time lags was set to $L = 25$ in a patient-specific paradigm. The patient with the greatest number of false positives per day when detecting all seizures has 2.5 false detections per day. The median number of false detections per day is 0.5. The median F_1 score across all patients is 0.95 (best between 100% sensitivity and detection of all but one seizure).

Table 3. Number of false positives per day for each patient to detect all seizures and to detect all but one seizure when all ($N = 18$) bipolar channels were used and the number of time lags was set to $L = 25$ in a patient-specific paradigm.

Patient	Recorded Seizures	False positives per day for 100% detection	False positives per day to detect all but one seizure	F_1 score
2	41	0	0	1
3	8	0.5	0.5	0.90
5	5	1.5	1.5	0.87
6	5	0.5	0	0.94
7	24	2.5	1	0.96
8	18	0	0	1

Figure 5 shows the number of false positives per day for a sensitivity of 95% when varying both the number of time lags (L) and the number of channels (N) in a patient-specific paradigm. The figure shows a decrease of false positives when the number of channels and/or the number of time lags increase.

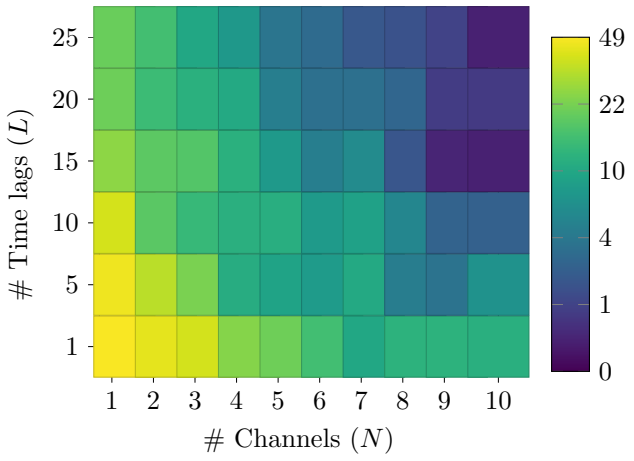


Figure 5. Median (over the different subjects) of seizure false positives per day when setting the sensitivity to 95% and varying the number of channels (N) and time lags (L) in a patient-specific paradigm. Color represents false positives per day. Dark blue represents a low number of false detections, yellow a high number.

Figure 6 shows how the performance of the algorithm varies in function of the number of channels (N) that are used for the classification task with $L = 25$ time lags in a patient-specific paradigm. The number of false positives per day decreases with increasing number of channels, reaching 0.5 false positives per day when using 10 channels.

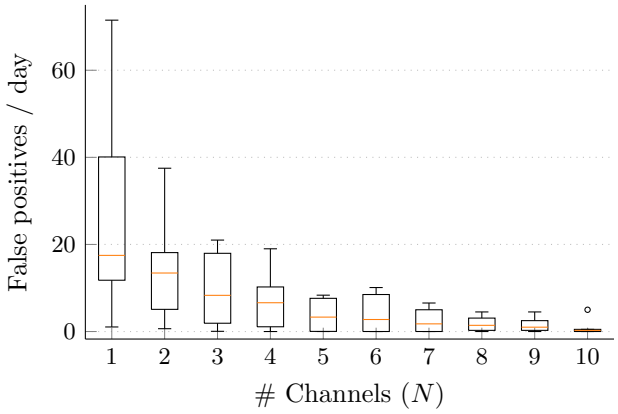


Figure 6. Boxplot (over the different subjects) of seizure false positives per day when setting the sensitivity to 95% and varying the number of channels (N) when using $L = 25$ time lags in a patient-specific paradigm.

Figure 7 illustrates the effect of the number of available training seizures on the performance of the algorithm. The algorithm was trained in a patient-specific fashion using all ($N = 18$) channels and $L = 25$ time lags. This graph shows results for three patients (patients 2, 7 and 8), the only ones with 10 or more seizures. Figure 7a illustrates that the number of available training seizures has very little influence on the performance of the algorithm. The spatio-temporal pattern in the seizure covariance matrix estimated from one seizure (Fig. 7b) is very similar to the seizure covariance matrix estimated from all seizures (Fig. 7c).

Figure 8 compares the performance of the algorithm when training in a patient-specific paradigm compared to a patient-independent paradigm. Both were generated using all ($N = 18$) channels and $L = 25$ time lags. The sensitivity versus false positive rate starts diverging for sensitivities higher than 60%. For these higher sensitivities the false positive rate per day is higher in the patient-independent paradigm than in the patient-specific one. For a sensitivity of 95% the patient independent-algorithm makes a median of 6 false positives per day while the patient-specific one makes 0.5 false positives per day.

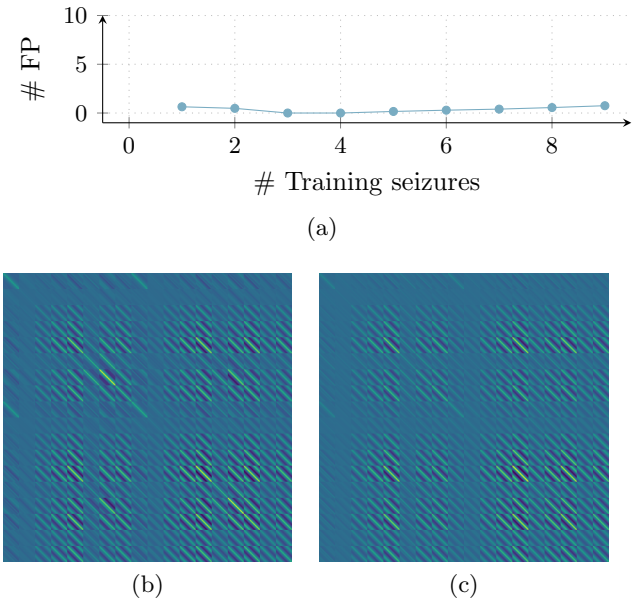


Figure 7. (a) Median number of false positives per day for a sensitivity of 95% when artificially limiting the number of training seizures, using all ($N = 18$) channels and $L = 25$ time lags in a patient-specific paradigm. – (b) Example of a seizure covariance matrix estimated from one seizure – (c) Example of a seizure covariance matrix estimated from all seizures in the same patient.

As a final validation, the patient-independent filter trained on the out-of-clinic recordings was evaluated on the validation dataset of patients recorded in hospital in order to test how well the pre-trained filter generalizes to other data sets. For a detection threshold that detects all 222 seizures in the dataset, there were two false detections in two different patients across the whole dataset. An example of a true detection and of a false detection are given in figure 9. Both detections occur while the patient is eating. The artifacts resulting from chewing are well suppressed in both examples. The seizure is correctly amplified. In the case of the false detection some noise with a seizure-like pattern triggers a false detection.

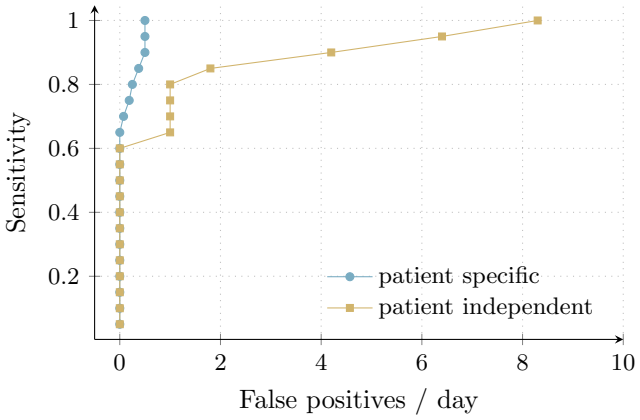


Figure 8. Median (over the different subjects) of seizure detection sensitivity in function of false positives per day when applying the data-driven filter with all ($N = 18$) channels and $L = 25$ time lags in a patient-specific paradigm (blue) and a patient-independent paradigm (ocher).

3.2. *Benchmarking performance and energy consumption*

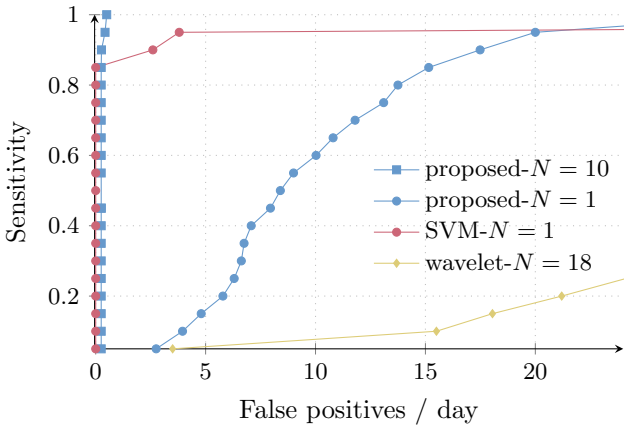
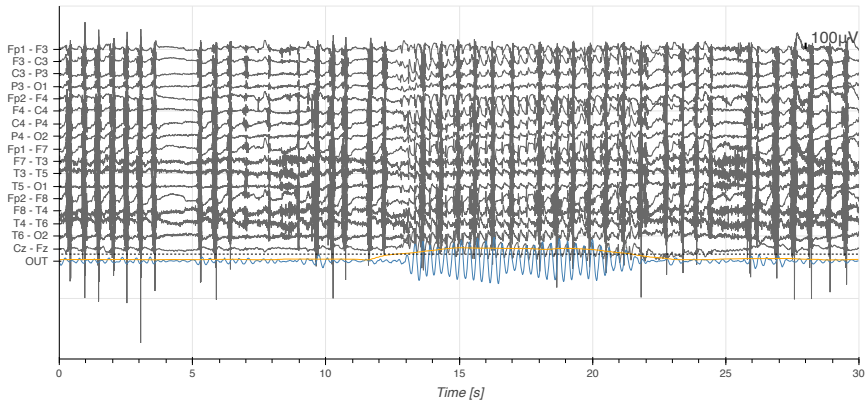


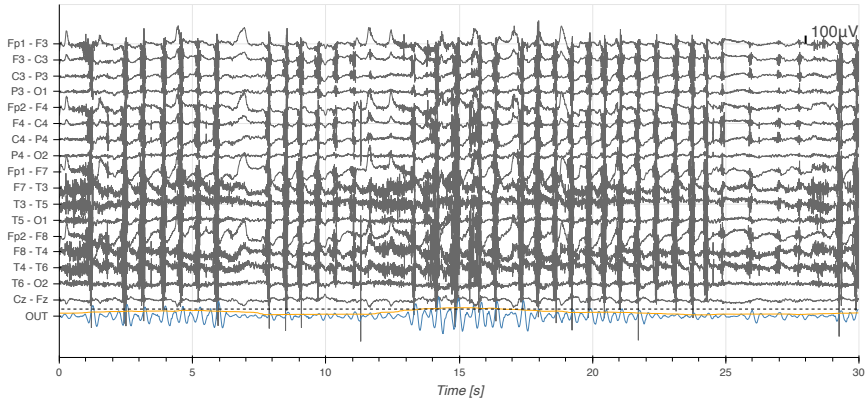
Figure 10. Comparison of the sensitivity in function of false positives per day in four different seizure detection algorithms in a patient-specific paradigm. SVM-$N = 1$ is the algorithm of Kjaer et al.[16] Wavelet-$N = 18$ is the algorithm of Xanthopoulos et al.[17]

Figure 10 shows a comparison of the performance of our proposed algorithm with state-of-the-art algorithms for patient-specific absence seizure detection. Four seizure detection algorithms are shown on the figure. The best performing classifier is a nonlinear SVM classifier on a single-lead EEG.[16] The best single-lead channel was investigated in the work of Kjaer et al. For their algorithm, they found F7−Fp1 to be the best channel. It has a median false positives per day across subjects of zero for sensitivities up to 85%. For higher sensitivities, the number of false positives increases rapidly. The classification method we propose ($N = 10, L = 25$) performs similarly with 0.25 false positives per day for a sensitivity of 85%. When the method we propose is restricted to one channel ($N = 1, L = 25$), it generates 15 false positives per day. The multi-channel ($N = 18$) wavelet-based algorithm does not perform well on this dataset.[17] It already has 15.5 false positives per day for a sensitivity of 10%. The number of false positives increases with higher sensitivities.

The memory usage of the method we propose corresponds to the data that are buffered and in the storage of the filter coefficients. The number of calculations needed to classify are the result of applying the data-driven filter and computing the RMS of the single-channel output. The memory usage of the nonlinear SVM classifier on a single-lead EEG[16] resides in a data buffer, the finite impulse response filter that is used to calculate some features and in the support vectors that need to be stored for classification of a new epoch. The SVM algorithm uses an average of 403 support vectors (full range: [17, 1659]). The number of calculations needed to classify a new epoch are the result of the feature extraction and classification using a radial basis function kernel. The memory usage of the wavelet-based method[17] lies in the data that are buffered. The number of calculations needed to classify a new epoch are the result of the continuous wavelet transform and the transformation to a single-channel output (variance of individual channels and mean over channels). The total number of bytes of memory as well as the number of operations (summations, multiplications, non-linear operations) are reported in table 4. The number of operations is reported as a normalized number of operations per sample. For the window-based algorithms, the number of operations on a window is divided by the number of samples in the window. For comparison purposes the sampling frequency of all algorithms was set to 200 Hz. The algorithm we propose uses 100 bytes of memory in the single channel and 1 kB of memory in the 10 channel configuration. The amount of memory required and the number of computations



(a)



(b)

Figure 9. An example of detected seizure (a) and of a false detection (b) in the validation dataset. The blue line is the single-channel output of the filter. The orange line is the running mean of the RMS of the single-channel output. The black dashed line is the detection threshold.

are linearly proportional to the number of channels and the number of time lags (see subsection 2.2.2). The SVM and the wavelet methods use respectively 12 kB and 8 kB of memory. The algorithm we propose executes 50 operations per sample period (25 summations and 25 multiplications) for one channel and 500 operations per sample period for 10 channels. The SVM uses a windowed approach and classifies windows of two seconds. The algorithm executes 6519 operations per sample period, of which 2 are non-linear (cosine and radial basis function). The wavelet algorithm classifies windows of one second. It executes 24012 operations per sample period. Details on the implementation of these algorithms that allows to calculate the number of operations is given in Appendix A.

Table 4. Memory consumption (bytes) and number of operations of four seizure detection algorithms. proposed-$N = 10$: our proposed method with $N = 10$ channels and $L = 25$ time lags, proposed-$N = 1$: our proposed method with one ($N = 1$) channel and $L = 25$ time lags, SVM-$N = 1$: a single-channel classifier based on a SVM,[16] wavelet-$N = 18$: a multichannel ($N = 18$) wavelet-based classifier.[17]

	Memory [bytes]	Operations per sample
proposed-$N = 10$	1000	500
proposed-$N = 1$	100	50
SVM-$N = 1$	11976	6519
wavelet-$N = 18$	8000	24012

4. Discussion

In this paper we propose an absence seizure detection algorithm specifically designed to run on a microcontroller or field-programmable gate arrays (FPGA). We collected data in real-life, out-of-clinic environments using a portable EEG amplifier. It has a sensitivity of 95% for 0.5 false positives per day when using 10 EEG channels and 25 time lags. The algorithm requires only 1000 bytes of memory and 500 operations per new sample (250 multiplications and 250 summations). This is well within the capabilities of typical microcontrollers that would be found in a wearable EEG device, such as ARM Cortex M4F.[10,12] The computational efficiency can be further increased by reducing the number of channels or the number of time lags (see Fig. 5 where the corresponding trade-off is visualized), the sampling rate, or by reducing the output sample rate of the spatio-temporal filter, thereby avoiding the need to compute a filter output sample for each input sample.

Low power requirements have several advantages for the patient. First, it allows the use of smaller size power supply, which is relevant considering that batteries are among the larger and heavier components in a wearable device. This facilitates further miniaturization and concealability of wearable EEG sensor devices. Second, requiring less power allows for a longer recording time, which in turn demands less manipulation effort from the patient, who would not need to replace batteries as often as in currently available devices. Furthermore, algorithms with a sufficiently low computational/memory footprint such that the processing can be on board of the device itself allow to bypass the need for a cloud-based processing, thereby eliminating an energy-intensive wireless transfer of the raw data.

A sensitive seizure detection algorithm is a valuable asset to a neurologist in order to have a precise idea of the number of seizures the patient is having. It is well known that patients with absence epilepsy have difficulties keeping an accurate seizure diary. The algorithm we propose could be used by the neurologist to considerably speed up the analysis of EEG recordings. For this use, the algorithm is set to a sensitivity of 100% and the neurologist is only given a handful of events (including some false positives) to review.

We investigated the effect of the number of channels on the performance of the algorithm. The performance of the algorithm greatly improved when more channels were available. Ten well-chosen channels were sufficient to reach 0.5 false positives per day for a sensitivity of 95% in all but one patient. Using three channels was associated with a median of eight false positives per day for a sensitivity of 95%.

We demonstrated that this method requires only a very small number of example seizure data to train a linear data-driven filter with good performance. Results from the algorithm barely varied when training with one seizure compared to training with more seizures. This was shown both in the performance of the seizure detection algorithm when trained on a low number of seizures and in the visual comparison of the seizure covariance matrices (figure 7). This is in contrast with other machine learning approaches such as neural networks or SVMs,[16,30,31] which require a large amount of training data before being able to classify correctly. The need for very few annotated samples to train our data-driven filter is a major advantage over complex machine learning methods as this allows neurologists to reduce the time they spend annotating EEG files. It also allows for the use of transfer learning techniques where a patient independent version of the filter is refined based on a small set of patient specific examples. While only a small number of seizures are required to train the algorithm, recording quality during these few examples influences the performance of the algorithm. Patient 5 had only five seizures during the recording period. All seizures were short and the recording was of relatively low quality (i.e. noise during the seizure). Patient 5 is the only patient to generate more than 1 false detection per day for a 95% sensitivity.

We compared the performance of our method in a patient-specific paradigm to the same method in a patient-independent paradigm. A patient-specific approach is associated with a lower number of false detections for the same sensitivity as it is able to better capture the specific characteristics of seizures and non-seizure data for that patient. This includes information on the temporal content of seizures and peak interferences as well as spatial information. A patient-independent paradigm allows the algorithm to be used on a patient for whom no previously annotated data are available. This considerably broad-

ens the pool of patients who could potentially benefit from this algorithm. However, the performance of the algorithm deteriorates in the patient-independent paradigm to six false detections per day. A transfer learning approach could be used where a patient-independent filter is used initially, and then updated to a patient-specific filter when more data from the patient become available. This is beyond the scope of this study.

We verified that our algorithm was generalizable to an independent dataset by applying the patient-independent version of the algorithm trained on patients recorded in a out-of-clinic environment to a dataset collected in the hospital. Results from these 17 in-hospital patients showed detection of 222 seizure events collected over more than 14 days of recording which resulted in only two false detections across all patients. The very low number of false positives illustrates that our recordings conducted in a hospital setting contained less artifacts than recordings collected in environments outside the clinic. False detections are usually triggered when the EEG contains high-amplitude noise and when the spatio-temporal characteristics of this noise happen to overlap with the spatio-temporal characteristics of seizures. As we showed in this study, this is a rare event. We note that the algorithm can be straightforwardly extended with a mechanism that estimates the event duration based on a thresholding of the filter output. In this case, it would be able to distinguish between long epileptiform events and short (typically interictal) epileptiform events. Furthermore, if short interictal events are not of interest to the neurologist, this mechanism can also be used to reduce the number of false positives due to short interfering events. However, since the distinction between ictal and interictal event duration is somewhat arbitrary and because the necessity of it is use-case specific, we did not take event duration into account in our analysis.

We compared our proposed algorithm with two state-of-the-art algorithms for absence seizure detection.[16,17] Kjaer et al.[16] designed an algorithm for wearable single-lead EEG where data analysis is performed off-line after the recording. The algorithm extracts many EEG features that differentiate an absence seizure from the rest of the EEG signal. A nonlinear SVM classifier is then used to classify two-second EEG epochs. The algorithm performs very well when it is provided with enough training data (more than five seizures). The SVM algorithm cannot be ported directly onto a microcontroller such as found on wearable EEG systems as the calculation of the EEG features, and classification are computationally heavy processes. The single-channel SVM classifier performed similarly to the multichannel version we propose. We also compared our method to a multichannel wavelet-based algorithm developed to detect absence seizures. The algorithm shares some underpinnings with the method we proposed as it is based on a filtering step, reduction from multichannel data to a single-channel output, and a threshold on the single-channel output. However, that algorithm does not use any spatial information on seizures, imposing an equal weight on all channels. That method does not have a data-driven mechanism to suppress peak interferences so that large amplitude artifacts would also trigger the thresholding mechanism leading to false positives. Another difference is that it uses wavelet-based filtering, which is computationally expensive. Xanthopoulos et al. report 4.8 false positives per day for a sensitivity of 80% on a dataset of six people with absence seizures with a total of 40 hours of recording.[17] The algorithm performed less well on our dataset. Xanthopoulos et al. identified that large amplitude artifacts trigger false detections with their method. In our ambulatory EEG recordings, many such high amplitude artifacts are present. These are suppressed by our proposed method by the spatial filtering.

The sample we used is small, though it appeared to be informative. We also focused on ca. 24h recordings, so that generalization to multiple-day recordings requires further testing. It must be emphasized that the algorithm focused on a single seizure type, which is appropriate in some but not all epilepsy disorders. Absence seizures are common in many forms of epilepsy across all ages and they are the hallmark seizure type in two epilepsy syndromes, namely childhood absence epilepsy and juvenile absence epilepsy. However, in many epilepsies, they are not the only seizure type occurring in a patient. In addition, other seizure types may have a less characteristic EEG signature and may involve less channels, making accurate detection more challenging.[32]

We believe this work is crucial to advance the development of wearable seizure detection systems for long-term monitoring of people with epilepsy outside

of the hospital. Additional work is required to further validate and test this algorithm and to actually integrate it in a wearable EEG device. In such a device, the location and distance between the electrodes will be constrained by wearability considerations.[33] Further work will need to be done to investigate how the distance between electrodes influences the algorithm performance. In addition, validation of this method in a larger patient cohort and during long term monitoring is needed before it can be applied in a clinical setting.

5. Conclusion

We have proposed a data-driven linear filtering method for absence seizure detection that is designed to run on a microcontroller for use in a wearable EEG-based seizure detection system. The algorithm aims to suppress peak interferers while enhancing seizure signal. It requires only a few annotated seizures to train the optimal filter. Our algorithm has been benchmarked against two state-of-the-art absence seizure detection algorithms (without computation constraints), and was found to perform almost on par with the best of these, while being much more efficient in terms of hardware memory and computational requirements.

Acknowledgments

Jonathan Dan is funded by VLAIO and Byteflies through a Baekeland grant. The authors acknowledge the financial support of the KU Leuven Research Council (C14/16/057) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 802895).

A conference precursor of this paper has been published as a 1-pager in.[34]

Appendix A Calculating the number of operations of each algorithm

When counting the operations of the algorithms in the work of Kjaer et al.[16] and Xanthopoulos et al.,[17] the following implementation of common algorithms were chosen:

- The fast Fourier transform (FFT) is implemented using the radix-2 Cooley-Tukey algorithm. The algorithm requires $5N\log_2(N)$ operations, where N is the number of samples in a window and is chosen as a power of two.[35]
- The continuous wavelet transform was implemented using FFTs. This implementation takes an FFT of the signal, an FFT of the wavelet, multiplies both signals and takes the inverse-FFT.[36] This process is repeated at every scale.

Bibliography

1. E. Beghi, G. Giussani, E. Nichols *et al.*, Global, regional, and national burden of epilepsy, 19902016: a systematic analysis for the global burden of disease study 2016, *The Lancet Neurology* **18**(4) (2019) 357 – 375.
2. Y. W. Cho, G. K. Motamedi and K. T. Kim, The clinical utility of non-invasive video-electroencephalographic monitoring has been diversifying, *Neurological Sciences* (aug 2019) 1–7.
3. C. E. Elger and C. Hoppe, Diagnostic challenges in epilepsy: seizure under-reporting and seizure detection, *The Lancet Neurology* **17**(3) (2018) 279 – 288.
4. T. Kobulashvili, J. Hfler, J. Dobesberger *et al.*, Current practices in long-term video-EEG monitoring services: A survey among partners of the E-PILEPSY pilot network of reference for refractory epilepsy and epilepsy surgery, *Seizure* **38** (2016) 38 – 45.
5. C. P. Panayiotopoulos, *Epileptic seizures and their classification*, *A Clinical Guide to Epileptic Syndromes and their Treatment* (Springer London, London, 2010), London, pp. 21–63.
6. H. Gastaut, Part i: Definitions, *Dictionary of Epilepsy*, (World Health Organization, Geneva, 1973), pp. 11–13.
7. M. J. Keilson, W. A. Hauser, J. P. Magrill and J. Tepperberg, Ambulatory cassette EEG in absence epilepsy, *Pediatr. Neurol.* **3**(5) (1987) 273–276.
8. I. Zibrandtsen, P. Kidmose, C. Christensen and T. Kjaer, Ear-EEG detects ictal and interictal abnormalities in focal and generalized epilepsy a comparison with scalp EEG monitoring, *Clinical Neurophysiology* **128**(12) (2017) 2454 – 2461.
9. F. Pinho, J. Cerqueira, J. Correia, N. Sousa and N. Dias, myBrain: a novel EEG embedded system for epilepsy monitoring, *Journal of Medical Engineering & Technology* **41**(7) (2017) 564–585.
10. D. Sopic, A. Aminifar and D. Atienza, e-Glass: A wearable system for real-time detection of epileptic seizures, *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, (Florence, Italy, 2018), pp. 1–5.
11. S. Debener, R. Emkes, M. De Vos and M. Bleichner,

Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear, *Sci Rep* **5** (Nov 2015) p. 16743.
12. S. Boeckx, W. van Paesschen, B. Bonte and J. Dan, Live demonstration: SeizeIT - a wearable multimodal epileptic seizure detection device, *2018 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, (Cleveland, Ohio, USA, 2018), pp. 1–1.
13. A. Bach Justesen, M. T. Foged, M. Fabricius *et al.*, Diagnostic yield of high-density versus low-density EEG: The effect of spatial sampling, timing and duration of recording, *Clin Neurophysiol* **130** (Nov 2019) 2060–2064.
14. C. Baumgartner, J. P. Koren and M. Rothmayer, Automatic computer-based detection of epileptic seizures, *Frontiers in Neurology* **9** (2018) p. 639.
15. U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis and J. S. Suri, Automated EEG analysis of epilepsy: A review, *Knowledge-Based Systems* **45** (2013) 147 – 165.
16. T. W. Kjaer, H. B. Sorensen, S. Groenborg, C. R. Pedersen and J. Duun-Henriksen, Detection of paroxysms in long-term, single-channel EEG-monitoring of patients with typical absence seizures, *IEEE Journal of Translational Engineering in Health and Medicine* **5** (2017).
17. P. Xanthopoulos, S. Rebennack, C. C. Liu *et al.*, A novel wavelet based algorithm for spike and wave detection in absence epilepsy *10th IEEE International Conference on Bioinformatics and Bioengineering 2010, BIBE 2010* , (IEEE, Philadelphia, Pennsylvania, USA, 2010), pp. 14–19.
18. H. Adeli, Z. Zhou and N. Dadmehr, Analysis of EEG records in an epileptic patient using wavelet transform, *Journal of Neuroscience Methods* **123**(1) (2003) 69 – 87.
19. V. Sakkalis, G. Giannakakis, C. Farmaki *et al.*, Absence seizure epilepsy detection using linear and nonlinear EEG analysis methods, *Conf Proc IEEE Eng Med Biol Soc* **2013** (2013) 6333–6336.
20. K. Zeng, J. Yan, Y. Wang *et al.*, Automatic detection of absence seizures with compressive sensing EEG, *Neurocomputing* **171** (2016) 497 – 502.
21. F. Manzouri, S. Heller, M. Dmpelmann, P. Woias and A. Schulze-Bonhage, A comparison of machine learning classifiers for energy-efficient implementation of seizure detection, *Frontiers in Systems Neuroscience* **12** (2018) p. 43.
22. A. Burrello, L. Cavigelli, K. Schindler, L. Benini and A. Rahimi, Laelaps: An energy-efficient seizure detection algorithm from long-term human iEEG recordings without false alarms, *Proceedings of the 2019 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, (IEEE, Florence, Italy, 2019), pp. 752 – 757.
23. M. Hugle, S. Heller, M. Watter *et al.*, Early seizure detection with an energy-efficient convolutional neural network on an implantable microcontroller, *arXiv e-prints* (June 2018).
24. J. Wouters, F. Kloosterman and A. Bertrand, Towards online spike sorting for high-density neural probes using discriminative template matching with suppression of interfering spikes, *Journal of Neural Engineering* **15**(5) (2018).
25. J. R. Tenney and T. A. Glauser, The current state of absence epilepsy: can we have your attention?, *Epilepsy Curr* **13** (May 2013) 135–140.
26. G. A. Worrell, T. D. Lagerlund and J. R. Buchhalter, Role and limitations of routine and ambulatory scalp electroencephalography in diagnosing and managing seizures, *Mayo Clinic Proceedings* **77**(9) (2002) 991 – 998.
27. B. D. V. Veen and K. M. Buckley, Beamforming: a versatile approach to spatial filtering, *IEEE ASSP Magazine* **5** (April 1988) 4–24.
28. J. Wouters, F. Kloosterman and A. Bertrand, A data-driven regularization approach for template matching in spike sorting with high-density neural probes, *2019 IEEE Engineering in Medicine and Biology Society Conference (EMBC)*, (Berlin, Germany, 2019).
29. S. Beniczky and P. Ryvlin, Standards for testing and clinical validation of seizure detection devices, *Epilepsia* **59**(S1) (2018) 9–13.
30. R. Hussein, H. Palangi, R. K. Ward and Z. J. Wang, Optimized deep neural network architecture for robust detection of epileptic seizures using EEG signals, *Clinical Neurophysiology* **130**(1) (2019) 25 – 37.
31. U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan and H. Adeli, Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals, *Computers in Biology and Medicine* **100** (2018) 270 – 278.
32. C. Baumgartner and J. P. Koren, Seizure detection using scalp-EEG, *Epilepsia* **59**(S1) (2018) 14–22.
33. A. M. Narayanan and A. Bertrand, Analysis of miniaturization effects and channel selection strategies for EEG sensor networks with application to auditory attention detection, *IEEE Transactions on Biomedical Engineering* (2019) 1–1.
34. J. Dan, B. Vandendriessche, W. V. Paesschen, D. Weckhuysen and A. Bertrand, Low-complexity data-driven seizure detection algorithm for home monitoring of patients with epilepsy using wearable eeg, *2019 IEEE Engineering in Medicine and Biology Society Conference (EMBC)*, (Berlin, Germany, 2019).
35. S. G. Johnson and M. Frigo, Implementing FFTs in practice, *Fast Fourier Transforms*, ed. C. S. Burrus (OpenStax CNX, Nov 2012).
36. M. J. Vrhel, C. Lee and M. A. Unser, Comparison of algorithms for the fast computation of the continuous wavelet transform, *Wavelet Applications in Signal and Image Processing IV* , **2825**, International Society for Optics and Photonics, (SPIE, 1996), pp. 422 – 431.