Electronic Sleep Stage Classifiers: A Survey and VLSI Design Methodology
Hossein Kassiri, Member, IEEE, Aditi Chemparathy, M. Tariqus Salam, Member, IEEE, Richard Boyce, Antoine Adamantidis, and Roman Genov, Senior Member, IEEE

Abstract—First, existing sleep stage classifier sensors and algorithms are reviewed and compared in terms of classification accuracy, level of automation, implementation complexity, invasiveness, and targeted application. Next, the implementation of a miniature microsystem for low-latency automatic sleep stage classification in rodents is presented. The classification algorithm uses one EMG (electromyogram) and two EEG (electroencephalogram) signals as inputs in order to detect REM (rapid eye movement) sleep, and is optimized for low complexity and low power consumption. It is implemented in an on-board low-power FPGA connected to a multi-channel neural recording IC, to achieve low-latency (order of 1 ms or less) classification. Off-line experimental results using pre-recorded signals from nine mice show REM detection sensitivity and specificity of 81.69% and 93.86%, respectively, with the maximum latency of 39 µs. The device is designed to be used in a non-disruptive closed-loop REM sleep suppression microsystem, for future studies of the effects of REM sleep deprivation on memory consolidation.

Index Terms—Alzheimer’s disease, implantable device, low latency, low-power FPGA, sleep, sleep classifier, sleep stage, VLSI implementation.

I. INTRODUCTION

EURODEGENERATIVE diseases affect millions of people worldwide. The most common type is the Alzheimer’s disease (AD), affecting approximately 5.9 million people only in the North America [1]. It is a progressive disorder that slowly erodes memory and thinking skills, and eventually results in inability to carry out simple tasks. Epidemiological studies have discovered that excessive REM (rapid eye movement) sleep is a potential risk factor for the AD [2].

As shown in Fig. 1, sleep is dominated by cyclic occurrences of SWS (slow-wave sleep) and REM stages. It is commonly accepted that during SWS, also known as non-REM (NREM) sleep, active consolidation of memory takes place by reactivation of the newly encoded memories, which are then integrated into the existing network of associated memories [3]. On the contrary, the REM sleep can possibly disrupt the memory consolidation process [4]. A recent study demonstrates that antidepressants drugs suppress REM sleep and no impairment of the consolidation of procedural memory [5] takes place. However, many patients have significant systemic side-effects and some are drug-resistant. These poor outcomes and adverse effects of the drugs motivate for an alternative treatment to supplement the conventional options for REM sleep suppression.

Electrical brain stimulation has been established as an effective drugs-alternative treatment option for a variety of neurological disorders [6], [7], [35]. For REM sleep, it has been shown that stimulating the lateral hypothalamus immediately after θ-oscillation (4–8 Hz) peaks and troughs, significantly increases the chance of REM suppression [9]. This necessitates the development of a sleep stage classifier that can detect REM stage, not only with a high accuracy, but also within a very short period of time (order of 1 ms or less), allowing for a timely stimulus to be delivered electrically [10], or optically [11].

Several mathematical algorithms have been proposed for sleep classification and are tested on off-line data using computer software [12], [13], [15], [16]. Benefiting from high-performance computational schemes, these algorithms often succeed to yield a high detection accuracy and a low false positive rate. However, they require high computational power, typically not available in portable/wearable devices. Also, using a data acquisition module that sends the recorded brain signals to a computer results in delays that are orders of magnitude longer than timing-requirements for a REM-suppressing responsive stimulation. On the other hand, some simpler methods are reported for clinical and commercial sleep classification.
In this paper, first, we review various types of electronic sleep stage classifiers and compare them in terms of classification accuracy, level of automation, implementation complexity, invasiveness, and targeted application. According to the review, multiple types of physiological signals must be used as sensory information to achieve the level of accuracy required for REM sleep detection. An EEG/EMG-based classification algorithm is then proposed with a computational complexity level that allows for its implementation on a low-power field-programmable gate array (FPGA), mounted on a miniature device worn by a small rodent. This yields a sub-ms classification latency which is shorter than the permitted 1% of the $\theta$-oscillation period (125–250 ms), to carry out effective REM-suppression [9].

The proposed algorithm architecture and performance are compared with two best-performing FPGA-compatible sleep classification algorithms reported in the literature [39], [41], using our intracerebral EEG (icEEG) and EMG recordings from nine mice. Rodents (mice) are chosen as an accurate and low-cost model of human sleep. All the three algorithms are based on extracting frequency components of the recorded EEG and EMG signals, followed by evaluating simple mathematical functions and thresholding. The REM sleep detection accuracy is evaluated, to illustrate the superiority of the proposed algorithm. Next, the proposed algorithm implementation in an FPGA that was assembled together with a multi-channel recording and stimulation ASIC (application-specific integrated circuit) on a small PCB (printed circuit board) is presented. It is optimized to reduce the use of hardware resources and power. The device is validated using off-line icEEG and EMG data, and its detection performance is compared to the state of the art.

Correspondingly, the rest of the paper is organized as follows. Section II reviews and compares existing sensory methods used for sleep stage classification. Section III reviews various PSG-based sleep stage classification algorithms implemented in software and hardware. Section IV describes the proposed algorithm and compares its MATLAB simulation results to the state of the art. Section V describes the electronic implementation of the proposed algorithm on a miniature device. Section VI presents the classification results for both software and hardware implementations of the algorithm and compares them with the state of the art.

II. REVIEW OF SLEEP-STAGE MONITORING SENSORS

Patient self sleep assessment report is a common method to obtain information in research and clinical studies. Although sleep diaries and questionnaires are one of the easiest and most affordable methods to collect data over a long period of time, the collected data are subjective and may not always be accurate in both healthy and sleep-disordered groups [17]–[19]. Moreover, such assessments are limited to sleep/awake results, and cannot provide any information on different sleep stages.

Using electronic sensors for the sleep stage classification removes subjectiveness of the results, and unlike the self assessment reports, can provide information regarding different sleep stages. Actigraphy (ACT), autonomic nervous system (ANS) activity, and polysomnography (PSG) are the most popular methods used in the development of both commercial and research-based electronic sleep stage classifiers [12], [13], [15], [16], [20]–[26], [37]–[40]. Fig. 2 shows the most common physiological and non-physiological sensors used in each method. In addition to the type of sensory signals used, these methods vary in the level of invasiveness, the classification accuracy, the level of automation, and the targeted application. Generally, using fewer and less-invasive sensors has the advantage of simpler and more comfortable data acquisition for the subject, which is of great importance in wellness applications. However, when sleep monitoring is performed for the purpose of detection or treatment of a neurological disorder, simplicity can be traded for higher classification accuracy.

Actigraphy is the most common method used in the commercial wellness devices that are typically designed to be used in conjunction with a smart phone [20]. Over several days, the acceleration of the extremities (typically wrist) is recorded using an accelerometer and is stored in the cell phone memory or a memory module embedded in the device. Later, the recorded data are fed to a computer-based classifier for sleep/wake classification. Most of the algorithms that have been proposed for actigraphy implementation, classify periods of low activity as sleep [21], [22]. As a result, they cannot cope with the problem of misclassifying low activity tasks, such as reading a book, lying on the bed, watching a movie or the case where the sensor band is not worn. This makes the use of actigraphy-based devices limited to simple wellness
Table I summarizes and compares the existing sleep stage monitoring sensory methods introduced in this section. As will be shown in the next sections, the algorithm suitable for implantable and wearable sleep stage classification proposed here is derived from the gold standard polysomnography sensory method and achieves a classification accuracy of over 93%.

III. REVIEW OF SLEEP CLASSIFICATION ALGORITHMS

Due to the high accuracy of PSG-based monitoring, the vast majority of reported sleep stage classification algorithms are based on this sensory method. In this section, first, we review various PSG-based algorithms implemented in software, and then describe hardware implementations of several existing classifiers that employ full or a partial PSG-based sensory system. The section ends with a discussion on challenges in hardware implementation of a PSG-based classification system, and a proposed viable solution.

A. Software-Implemented Algorithms

Diagnostics of sleep disorders requires extracting information about one or a few specific sleep stages, which might include duration of the stage, and EEG/EMG time- and frequency-domain activity during that stage [12], [14]. As a result, the goal of the majority of sleep classification algorithms reported in the literature is to detect all the sleep stages that happen during a full sleep cycle with the highest possible accuracy [12]–[14], [16]. This level of specificity and accuracy requires either an increment in the number of input sensory signals, or a high level of computational complexity of the algorithm used, or both.

In [12], a time-frequency image (TFI) of EEG signals are used to perform sleep stage classification. The TFIs are segmented based on the frequency bands in order to extract features that are later used as inputs of a multi-class least squares support vector machine (MC-LS-SVM) with different kernel functions. Different kernels were compared in terms of their accuracy in sleep stage classification. The reported SVM can classify all 5 sleep stages with an overall accuracy of 88.47%.

In [13], wavelet packet coefficients and artificial neural networks (ANN) are applied to a pre-recorded EEG dataset to conduct sleep stage classification. As it is reported in the paper, REM and stage 1 of NREM sleep are indistinguishable using EEG signals only. As a result, the method that has been used can only classify awake, stage 1 + REM, stage 2 and the slow wave stage, but with a high average accuracy of 93%, thanks to the sophisticated ANN used.

In [14], an 18-channel polygraph is used for sleep classification that includes five EEG channels, EOG for REM detection, tonic chin and diaphragmatic EMGs, electrocardiogram (ECG), body movement detection of upper and lower limbs using piezo-electric crystal transducers, abdominal respiratory movements using a mercury strain gauge, and nostrils airflow, by means of a thermistor. A neuro-fuzzy classifier (NFC) is employed to classify sleep/wake stages and stages was trained, validated and tested on an offline dataset of 14 healthy infants of ages 6 months old and onward. After training with 7 datasets and validating

Fig. 3. (a) Placement of different head sensors used for polysomnography (PSG). (b) Approximate average signal power in the $\delta$ (cortex) and $\theta$ (hippocampus) EEG bands, and the EMG band during different sleep stages.
with another 2 datasets, the NFC is tested on 5 datasets and can classify all sleep stages with an overall classification accuracy of 84%.

In [16], an SVM with a Gaussian radial basis kernel is used to classify between NREM and paradoxical sleep. ECoG and EMG signals are recorded from 6 rats and several time- and frequency-domain features are extracted from the recordings. Thanks to the computationally-sophisticated SVM used, an average 96% classification accuracy is achieved.

Besides the classification algorithms described above, there is a second group of algorithms mostly aimed at detection of a specific stage and are optimized for both simplicity and performance at the same time. These systems typically rely on simple mathematical functions applied on a feature that is extracted from one or a few physiologically recording types to detect the target sleep stage. Two examples of such systems are reported in [39], [41], in which both algorithms are aimed at classification between REM and NREM sleep without any further information on different NREM stages. Both works use simple mathematical computations for detection, and report high detection accuracy of above 88%. The details of each algorithm are described in Section IV.

Since a hardware implementation (for low latency) of a REM sleep detection algorithm is the main goal of our work, simplicity of the algorithm is of an equal importance to its detection accuracy. As a result, the algorithm proposed in this paper falls within the second group described above, which are aimed at performing a high-accuracy classification of a subset of sleep stages using a simple hardware-implementable method.

B. Hardware-Implemented Algorithms

There are only a small number of hardware implementations of sleep stage classification algorithms reported in the literature. In [32], a PSG-based fuzzy neuro generalized learning vector quantization (FNGLVQ) is used on EEG and ECG signals to perform sleep stage classification. The algorithm is implemented on a high-end Xilinx Spartan-3AN XC700AN FPGA with an average power consumption of 40 mW and yields a 68.8% classification accuracy and a 790 ms delay. The delay is mainly due to the fact that the main part of the algorithm is implemented in software (due to limited resources available on the FPGA) and the FPGA is required to send the signals to a computer for classification after performing some basic arithmetic operations. Although four different sleep stages are classified, neither the power consumption nor the delay meet the requirements for a battery-powered implantable or wearable device for sleep classification.

To implement a classification algorithm within the power and resource limits of a miniature device, the majority of reported systems only use a subset of PSG sensory signals. However, the system performance in sleep stage classification highly depends on the selected subset of sensory signal types. In [33], a portable sleep stage classifier is reported that works with signals from one lead of ECG (three electrodes) as inputs. The device is capable of differentiating between sleep and awake stage using a random forest algorithm implemented on an ATME7 microcontroller. The reported average classification delay is 20 seconds. Another ECG-based multi-lead wearable system for monitoring patient sleepiness (i.e., asleep versus awake only) is presented in [37]. The system is comprised of three ECG gel electrodes, EMG and EOG electrodes, inductive belt sensor, three electronic modules, and a NiMH battery. It is tested on multiple users and yields a 85.3% classification accuracy.

When the goal is to detect more than one non-awake sleep stages, EEG must be included in the selected subset of PSG sensory signals. As shown in Fig. 3(b), the  (4–8 Hz) oscillations from the hippocampus EEG are the most prevalent during the REM sleep and awake stages, and the  (0.5–4 Hz) oscillations from the cortex EEG are found during NREM sleep. In addition, neck muscle EMG high-frequency oscillations (100–200 Hz) can be used to distinguish between the REM and awake stages. This motivates for an algorithm that classifies sleep stages based on the extracted frequency components of the recorded EEG and EMG signal in three different bands.

This algorithm must be compact enough to be implemented on a low-weight and low-power device with a system architecture such as the one depicted in Fig. 4. As shown, the device has three low-noise sensory channels for physiological signal recording, and a digital back-end processor for algorithm implementation. To make the device untethered, it is equipped with a wireless module that transmits data for storage or further processing, and receives commands for system reconfiguration. Embedding responsive stimulation capability into such a device

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>SUMMARY OF SENSORY METHODS FOR SLEEP STAGE MONITORING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stages Classified</td>
<td>Actigraphy (ACT)</td>
</tr>
<tr>
<td>Sensory Signal Types</td>
<td>Awake/Sleep [20–24]</td>
</tr>
<tr>
<td>Result Interpretation</td>
<td>Motion &amp; Acceleration</td>
</tr>
<tr>
<td>Classification Accuracy</td>
<td>&lt;40%</td>
</tr>
<tr>
<td>Invasiveness</td>
<td>No</td>
</tr>
<tr>
<td>Implementation</td>
<td>Wireless Headset</td>
</tr>
</tbody>
</table>
enables REM sleep suppression for the purpose of improving the memory consolidation process, and realizes a promising therapeutic option for the treatment of neurological disorders such as Alzheimer’s disease. While the full system is shown in Fig. 4, this work focuses on the wearable classifier only.

IV. ALGORITHM COMPARATIVE STUDY

As described above [and illustrated in Fig. 3(b)], three filtered physiological signals provide enough information to distinguish between three stages of awake, REM, and non-REM sleep. Next, a low-complexity sleep classification algorithm is proposed. It is first implemented in MATLAB together with two best-performing REM-detection algorithms from the literature [39], [41]. Their performance is compared using an off-line sleep dataset as the input. The three algorithms selection was based on the following criteria, besides performance: the use of rodent physiological signals as inputs (for fair comparison) and low computational complexity (all perform classification by evaluating simple mathematical functions on extracted signal spectrum information as needed to be implementable on a low-power FPGA).

A. Three Selected Algorithms

1) Algorithm A: The first algorithm, shown in Fig. 5, was originally proposed in [39], and uses a locomotion sensor, EEG, and EMG signals. The power spectrum of the EEG signal is analyzed using FFT with a Hanning window to extract amplitude of the delta and theta waves. Similarly the \( \theta \)-band EEG activity during awake and REM stages makes them indistinguishable using EEG only. As a result, EMG signals are also used, and comparing the magnitude of the delta wave, the ratio of \( \theta / (\delta + \theta) \), and the integral of the EMG signal to different threshold values, the system classifies NREM, REM, and awake stages. The locomotor signal is used to distinguish between active and quiet awake stages.

2) Algorithm B: This algorithm which was first introduced in [41] is shown in Fig. 6. It compares the filtered EMG signal against a threshold value to detect the active wake stage, characterized by high EMG activity. To classify between the remaining stages, it filters the EEG signal to obtain \( \delta, \theta, \alpha, \beta, \) and \( \gamma \) band components. Two ratios are calculated and compared with their respective thresholds to classify NREM, REM, and quiet wake stages.

3) Algorithm C (Proposed): In the third algorithm, proposed in this work and shown in Fig. 7, three signals are acquired: one EEG signal from the hippocampus, one EEG signal from the cortex, and one EMG signal from a neck muscle. The EEGs from the cortex and hippocampus are filtered to obtain the signal components in the \( \delta \) and \( \theta \) bands, respectively. The EMG signal is also filtered between 100 Hz to 200 Hz. All three filters are fourth-order Chebyshev filters with 0.5 dB peak-to-peak passband ripple. The filtered signals are passed through a window averaging block to obtain the mean amplitude of each signal to avoid instantaneous misclassifications. A ratio of the two filtered EEG signals \( \theta / \delta \) is then taken and compared with a patient-specific threshold value to distinguish between NREM and REM sleep. The EMG signal is also compared against another threshold value for awake stage detection.

B. Data Collection

To compare the performance of the three algorithms, they were initially implemented and tested in MATLAB using the
EEG and EMG signals that were collected from nine mice, and labeled by a trained sleep neuroscientist as described below:

1) Animals: Nine male C57 mice (from Charles River Lab, Quebec, Canada) were used in the experiments. The entire experiment was reviewed and approved by the animal care committee of the Douglas Health Institute (Montreal, Canada) according to the Canadian Guidelines for Animal Care.

2) Electrodes: Tetrode electrodes were composed of 17.5 μm diameter platinum-iridium wire (platinum:iridium 90:10%) with Teflon and VG bond coating (California Fine Wire company, Grover Beach, CA, USA), for insulation and heat-induced annealing, respectively. Epoxy served to insulate any exposed wire at the connection point with the electrode interface board. The annealed tip of the tetrode was cut using sharp scissors immediately prior to surgery to give an impedance of 1 MΩ.

3) Surgery: At 18 weeks age, injected mice were anesthetized with isoflurane (5% induction, 0.5–2% maintenance). The skull was completely cleared of all connective tissue and thoroughly dried using alcohol. For hippocampal electrode placement, a hole was drilled through the skull above the dorsal hippocampus (AP, −2.45; ML, +1.8), with the end target being the dorsal hippocampal CA1 pyramidal cell layer. An EEG screw was placed in the skull above the hippocampus in the contralateral hemisphere (AP, −2.3; ML, −1.35). An EMG electrode consisting of stranded tungsten wires (A-M Systems, WA, USA) inserted into the neck musculature was used to record postural tone. Screws placed in the bone above the frontal cortex and cerebellum served as ground and reference, respectively. Following surgery, mice were allowed to recover undisturbed for at least 1 week. Once mice had sufficiently recovered from surgery they were briefly anesthetized with isoflurane (5% induction, 2% maintenance) and a custom-built headstage pre-amplifier tether (Neuralynx, Inc., Boseman, MT, USA) was attached to a connector on the top of the implanted electrode interface board and secured with several small drops of epoxy.

4) Scoring: Recordings began only after mice were habituated to being chronically tethered. All recorded signals from implanted electrodes were amplified by the headstage pre-amplifier tether before being sampled and digitized at 16000 Hz. For scoring of behavioral state recorded data from hippocampal LFP, EEG and EMG electrodes, each raw data file was first imported into MATLAB (MathWorks, Natick, MA, USA) and downsampled to 1000 Hz. Data was then plotted and the vigilance state was manually scored in 5 s epochs. Scoring was based on visual characteristics of the hippocampal LFP, EEG and EMG data as well as fast Fourier transform analysis of each epoch scored. Wakefulness was defined by a de-synchronized low-amplitude EEG and hippocampal LFP and tonic EMG activity with periods of movement-associated bursts of EMG activity. NREM sleep was defined as synchronized, high amplitude, low-frequency (0.5–4 Hz) EEG and hippocampal LFP activity that was accompanied by reduced EMG activity relative to that observed during wakefulness. REM sleep was defined as having a prominent theta rhythm (4–8 Hz) and an absence of tonic muscle activity.

C. Comparative Study Simulation Results

The outputs of the three algorithms implemented in MATLAB using the data collected as described above are shown in Fig. 8. The first and second plots depict the two EEG signals from the hippocampus and the cortex, respectively. The third plot shows the filtered EMG signal from the neck. The fourth plot is the hypnogram, which is the result of manual sleep-stage scoring by a trained sleep neuroscientist as described in Section IV-B4 and is the reference against the output of various algorithms. The remaining six plots show the outputs of the three algorithms, A, B, C, and their corresponding average cumulative error calculated from point-to-point comparison with the hypnogram.

For each new animal, the algorithm is reconfigured to yield the best overall performance for sleep stage classification. To do so, the first 10% of each dataset is used to optimize the threshold values of the algorithm (for two EEG and one EMG signals), and the remaining of the dataset is used to evaluate the performance. The threshold values for each animal are chosen to maximize the overall accuracy, sensitivity and specificity.

The REM detection sensitivity, accuracy, and specificity are defined to evaluate each algorithm’s performance, and compare it to the state of the art as follows: True positives (TP): REM stage is correctly classified as REM. False positives (FP): NREM or awake stages are classified as REM. True negatives (TN): NREM or awake stages are correctly classified as NREM and awake, respectively. False negatives (FN): REM stage is classified as NREM or awake. Sensitivity: the ratio of TP to TN; Specificity: the ratio of TN to TP + FN. Specificity: the ratio of TN to TN + FP. The third algorithm, algorithm C, was chosen for hardware implementation because of its low complexity and better performance in terms of REM sleep stage detection average error.

The three algorithms have been compared in terms of their REM sleep detection sensitivity, specificity, and accuracy with each other as well as with all the algorithms described in Section III-A, with the comparison results summarized in Table II. In addition to the classification results reported in [39], [41], the algorithms from these works, which were...
described in Section IV, were also simulated in MATLAB with our animal dataset for a more precise comparison. Based on this comparison, the proposed algorithm, Algorithm C, yields superior classification performance compared to the algorithms reported in the literature.

V. ELECTRONIC IMPLEMENTATION OF ALGORITHM C

As mentioned in Section I, to stimulate efficiently for REM sleep suppression, the REM sleep detection latency has to be less than 1% of the $\theta$-oscillation period (125–250 ms). This motivates for an FPGA-based digital implementation to avoid long delays caused by a data acquisition module and a computer that are required for software-based implementa-

A. Hardware Implementation

The neural interface design is based on an integrated circuit (IC), that was first reported in [34]. It is used in this system and is configured to perform three-channel signal recording. The rest of the channels are shut off in the case of the application discussed in this paper to keep the focus (but can be used for simultaneous monitoring of other signals and for neurostimulation). The IC was fabricated in a 0.35 $\mu$m CMOS technology. Four recording channels and one stimulating channel occupy an area of $0.4 \times 0.4$ mm$^2$ as shown in Fig. 9 (left). The IC is directly wire-bonded to a PCB shown in Fig. 9 (right), and is protected by epoxy. The PCB is sized at $22 \times 30$ mm$^2$ and, besides the chip, houses several other components such as a low-power ACTEL FPGA, voltage regulators, and a crystal oscillator.

Fig. 10 shows the system-level block diagram of the multi-channel recording module. The system performs bipolar recording and uses Omnetics connectors to interface with the microelectrodes. As shown, the outputs of the recording amplifiers are multiplexed and sent to an ADC module that converts each recorded signal to an 8-bit digital output. The digital signals are then fed to the on-board FPGA for signal processing. The FPGA also controls a DAC (digital-to-analog converter) module that generates multiple bias voltages for the chip and configures stimulation parameters.

Additionally a separate bluetooth transceiver PCB of the same size is stacked with this system to provide a wireless link that communicates the diagnostic data and classification results to a computer [35]. The wireless link can also be utilized to receive commands to configure the system mode of operation. The system also has a current-mode neurostimulator capable of delivering biphasic charge-balanced pulses within the standard safety limit [36]. These functionalities are not the focus of this paper and are not discussed here further.

B. Algorithm Implementation

This paper focuses on the algorithm implementation. Hardware implementation of the algorithm minimizes classification time and makes the system capable of low-latency recording, detection and, if needed, closed-loop stimulation.

For a rodent-wearable device, the trade-off is the limited resources (logic elements) available on a low-power FPGA compared to the computational power of a computer, or even a high-end FPGA typically used in bench-top systems. For filter implementation, FIR topology was chosen over IIR due to the high sensitivity of IIR coefficients which makes the algorithm performance susceptible to noise. 64-tap FIR filters were chosen for FPGA implementation, as MATLAB simulation results showed that this is the minimum number of taps required
TABLE II
COMPARISON WITH THE EXISTING SOFTWARE-BASED SLEEP CLASSIFICATION METHODS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Source</td>
<td>As reported by authors</td>
<td>As simulated in MATLAB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stages Classified</td>
<td>W. REM S1, S2</td>
<td>W. REM S1, S2</td>
<td>W. REM, NREM</td>
<td>W. REM, NREM</td>
<td>Active/Quiet W. REM, NREM</td>
<td>W. REM, NREM</td>
<td>W. REM, NREM</td>
<td>W. REM, NREM</td>
<td></td>
</tr>
<tr>
<td>Signals Used</td>
<td>EEG</td>
<td>EEG</td>
<td>PSG</td>
<td>ECoG, EMG</td>
<td>EEG, EMG, locomotion</td>
<td>EEG, EMG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>TFL, MC-LS-SVM</td>
<td>ANN</td>
<td>Neuro fuzzy system</td>
<td>SVM</td>
<td>Motion sensor FTT &amp; thresholding</td>
<td>Filtering &amp; thresholding</td>
<td>Motion sensor FTT &amp; thresholding</td>
<td>Multi-band FTT &amp; thresholding</td>
<td>Multi-band FTT &amp; thresholding</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>88.47</td>
<td>93.0</td>
<td>83.9</td>
<td>&gt;96</td>
<td>90.0-90.9</td>
<td>87.9</td>
<td>84.34</td>
<td>82.54</td>
<td>93.1</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>N/R</td>
<td>84.2</td>
<td>N/R</td>
<td>N/R</td>
<td>N/R</td>
<td>N/R</td>
<td>71.02</td>
<td>66.75</td>
<td>82.9</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>N/R</td>
<td>94.4</td>
<td>N/R</td>
<td>N/R</td>
<td>N/R</td>
<td>N/R</td>
<td>96.4</td>
<td>84.3</td>
<td>93.9</td>
</tr>
<tr>
<td>Dataset</td>
<td>8 adults</td>
<td>7 adults</td>
<td>5 infants</td>
<td>6 rats</td>
<td>10 rats, 23 mice</td>
<td>14 rats</td>
<td>9 mice</td>
<td>9 mice</td>
<td>9 mice</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>N/R</td>
<td>N/R</td>
<td>N/R</td>
<td>N/R</td>
<td>N/R</td>
<td>N/R</td>
<td>2.73</td>
<td>2.86</td>
<td>2.43</td>
</tr>
</tbody>
</table>

N/A: Not applicable
N/R: Not reported

1 Data for training, validation and test datasets (14 infants in total) was reported. In this table only test data is shown.

2 TR: Transition-to-REM

Fig. 9. (a) The 30 mm × 22 mm PCB assembly of FPGA and the neural recording IC. (b) Envisioned implantation of the sleep classification device on a mouse head.

Fig. 10. Simplified hardware implementation of the small-scale VLSI-based sleep-stage classifier.

fit the algorithm in the FPGA, a single filter (with variable coefficients) was time-shared among three input channels. For this reason the clock frequency of the FPGA (40 MHz) was chosen to be much higher than the sampling rate of the input signals. Additionally, to further improve filtering performance, and consequently detection accuracy, every input channel was filtered twice using the same filter.

A block diagram of the hardware implementation of algorithm C is shown in Fig. 11, depicting how the input channels are sharing and reusing the FIR filter. The first multiplexer is used to time-share the FIR filters among the three channels, and the second one is used to select whether to filter a new signal or refilter the previous one for better band selectivity. The control block sends appropriate commands to control the logic and timing of each block. Following the band-pass filter, the three channels are averaged over a moving window with
Fig. 11. FPGA VLSI architecture implementation of the presented sleep-stage classification algorithm.

Fig. 12. Effect of the averaging window size on the sensitivity, specificity and accuracy of the REM sleep detection.

an adjustable size. As this requires significantly less resources compared to FIR filtering, a separate window averaging block is implemented for each channel, which results in lower classification latency (due to no time sharing), and also allows for a different window size for each channel.

VI. PROPOSED ALGORITHM RESULTS

A. Simulation Results

To further enhance the performance of the proposed algorithm, the window averaging block was optimized for REM sleep detection. The window length used is a trade-off between sensitivity and specificity. A longer window allows for a larger portion of the signal to be analyzed, increasing the specificity of the detection, while a shorter window increases the sensitivity of the detection. In Fig. 12, the relationship between window size, accuracy, sensitivity, and specificity is shown. The average value of the three measures was used for overall performance optimization, which resulted in a window size of 8 seconds.

B. Experimental Results

ASIC experimental measurement results show that each recording channel has a programmable gain of 54–72 dB, and programmable bandwidth with maximum of 1–5 kHz. The integrated input-referred noise is measured to be 7.99 $\mu$V$_{rms}$ and the total area and power consumption of each recording channel are 0.02 mm$^2$ and 12.9 $\mu$W, respectively.

The algorithm was implemented on an Actel ProASIC3 FPGA and tested with data from 9 different mice. The FPGA has an average total power consumption of 3.6 mW, but only has 3000 logic elements for the algorithm implementation. With the algorithm optimizations done for hardware implementation, it uses 2671 logic elements (89.03% of total), four on-chip RAMs to store 64 signal samples (one for the FIR filter and three for the averaging filters) and a flashROM to store filtering coefficients.

Fig. 13 shows a typical sample output of the FPGA implementation for a 9-minute recording, compared with the reference hypnogram and software implementation results. The first plot is the hypnogram, which was described previously and is the reference against the output of classification algorithm. The second and third plots depict the classification output of the software and hardware implementation of the proposed algorithm, respectively. The fourth and fifth plot show the instantaneous and cumulative average point-to-point errors of the FPGA implementation results compared with the hypnogram, respectively. The system needs 1562 clock cycles for every sample to generate an output which translates into 39 $\mu$s latency using a 40 MHz FPGA clock.
TABLE III

<table>
<thead>
<tr>
<th>Reference</th>
<th>Stages classified</th>
<th>Signal(s) used</th>
<th>Method</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Computation time (ms)</th>
<th>Real-time</th>
<th>Order of filter</th>
<th>Dataset</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>W, Non-wake</td>
<td>ECG</td>
<td>Random forest</td>
<td>N/R</td>
<td>68.8</td>
<td>94.2</td>
<td>20k</td>
<td>Yes</td>
<td>3</td>
<td>10 adults</td>
<td>N/R</td>
</tr>
<tr>
<td></td>
<td>W, Sleep</td>
<td>ECG</td>
<td>FNGLVQ</td>
<td>77.8-89.0</td>
<td>91.9</td>
<td>3.75^1</td>
<td>790</td>
<td>No</td>
<td>N/A</td>
<td>10 adults</td>
<td>3.13</td>
</tr>
<tr>
<td></td>
<td>W, REM, NREM</td>
<td>ECG, Resp.</td>
<td>FPT, ANN</td>
<td>81.66</td>
<td>93.86</td>
<td>0.039</td>
<td>Filtering+ thresholding</td>
<td>N/A</td>
<td>64</td>
<td>9 mice</td>
<td></td>
</tr>
</tbody>
</table>

N/A: Not applicable
N/R: Not reported

1 Estimated value. Algorithm is not implemented due to high power consumption.

Table III shows the accuracy, sensitivity and specificity of the presented implementation compared to the other hardware-based implementations described in Section III-B. In addition to classification performance, they are also compared in terms of the sensory signals used, the classification method, and their classification latency. The presented system yields the smallest latency, while exhibiting a comparable level of accuracy, sensitivity and specificity. It also has the smallest size as needed for a rodent-wearable device.

C. Discussion

Despite superior performance of the hardware implementation compared to the state of the art (as summarized in Table III), the FPGA implementation shows a 11% degradation in the sleep stage classification accuracy compared to MATLAB implementation. Generally, this difference is caused due to the limited resources (power and logic elements) available on the FPGA compared to the MATLAB implementation which has a virtually-unlimited computational power. In particular, we believe that the following are the major reasons for the performance degradation.

The first reason is the non-ideal band-pass filters compared to what is used in software implementation. As mentioned in the paper, for the hardware implementation, 64-tap FIR filters are used. Although higher number of taps would have allowed for higher filtering selectivity, and consequently better overall sleep classification, we could not increase it more than 64 due to the resource constraints of the FPGA (the 64-tap FIR filter is already the most resource-hungry component of this algorithm).

Our simulation results show an increase of 3% in accuracy when the filter length is increased from 64 to 256 taps.

The second reason which is also related to the filters is from loss of some samples of the EMG data. As the band-pass FIR filter is shared among the three channels of EEG1 (0.5–4 Hz), EEG2 (4–8 Hz), and EMG (100–200 Hz), the sampling frequency of the EMG channel (800 S/sec) is significantly higher than the other two (800 S/sec). As a result, when the filter is time-shared among the channels, three out of four samples of the EMG signal are skipped to ensure synchrony of operation among the channels. This loss of data which degrades the classification performance, can only be avoided if a large register is used as a FIFO for the EMG channel, which is not possible in this work due to resource limitations of the FPGA. Our simulation results show an increase of 6% in accuracy when dedicated filters with optimized sampling frequencies are used for each channel.

The third reason is due to the error that is caused by converting the floating point output of the digital divider to a fixed point number. The last source of error is simply the quantization noise caused by the accuracy of 8 bits (roughly translates to around 12 mV) for the threshold values. Our simulation results show an increase of 1.5% in classification accuracy when the quantizer resolution is increased from 8 to 12 bits.

VII. Conclusion

A low-latency, small-form-factor microsystem for sleep stage classification and REM sleep detection is presented. Three EEG and EMG signals are utilized to classify REM, NREM, and awake stages. The FPGA implementation is optimized to reduce complexity and power consumption while maximizing REM sleep detection performance. Experimental results show a REM detection sensitivity and specificity of 81.69% and 93.86%, respectively. A low latency of 39 µs has been achieved, which is a critically important design requirement for a closed-loop sleep control system. Such a system can be used for studies aimed at determining the effects of REM sleep suppression on memory consolidation.

REFERENCES


Dr. Antoine Adamantidis received pre- and postdoctoral education at the University of Liege, Liege, Belgium, and Stanford Medical School, Stanford, CA, USA. Currently, he is an Assistant Professor in the Department of Neurology, University of Bern, Bern, Switzerland, and holds a joined appointment in the Department of Clinical Research. He is the Codirector for the Zentrum for Experimentale Neurologie (ZEN labs) at the Inselspital in Bern. Also, he is an Adjunct Professor in the Department of Psychiatry at McGill University, Montreal, QC, Canada. His research interests aim at understanding the wiring diagram and the dynamics of sleep-wake circuits in the mammalian brain. His research objectives aim at investigating the wiring, firing dynamics, and plasticity of the neural circuits regulating brain states in normal and pathological states using in vitro and in vivo optogenetics—a technology that he and his colleagues pioneered at Stanford University—combined to genetics and electrophysiological methods. In his recent work, his laboratory identified a rapid-eye movement (REM) sleep circuit in the hypothalamus that control switch and duration of REM sleep states in mammals.

Dr. Adamantidis has received several awards, including the R. Broughton Young Investigator Award (Canadian Sleep Society), a Canadian Research Chair in Neural circuits and Optogenetics, a NIH Pathway to Independence (PI) Award-K99/R00 (USA), NARSAD, and Sleep Research Society Young Investigator Award (USA).

Roman Genov (S’96–M’02–SM’11) received the B.S. degree in electrical engineering from the Rochester Institute of Technology, NY, USA, in 1996, and the M.S.E. and Ph.D. degrees in electrical and computer engineering from Johns Hopkins University, Baltimore, MD, USA, in 1998 and 2003, respectively.

Currently, he is a Professor in the Department of Electrical and Computer Engineering at the University of Toronto, Toronto, ON, Canada, where he is a member of the Electronics Group and Biomedical Engineering Group and the Director of Intelligent Sensory Microsystems Laboratory. His research interests are primarily in analog integrated circuits and systems for energy-constrained biological, medical, and consumer sensory applications, such as implantable, wearable or disposable sensory Microsystems, energy-efficient sensory signal processors and wireless sensors, including brain-chip interfaces, neuro-stimulators, image sensors, optical and electrochemical DNA microarrays, and other biosensors.

Dr. Genov was a corecipient of Best Paper Award of the IEEE Biomedical Circuits and Systems Conference, Best Student Paper Award of the IEEE International Symposium on Circuits and Systems, Best Paper Award of the IEEE Circuits and Systems Society Sensory Systems Technical Committee, Brian L. Barge Award for Excellence in Microsystems Integration, MEMSCAP Microsystems Design Award, DALSA Corporation Award for Excellence in Microsystems Innovation, and Canadian Institutes of Health Research Next Generation Award. He was a Technical Program Cochair at the IEEE Biomedical Circuits and Systems Conference. He was an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–II: EXPRESS BRIEFS and IEEE SIGNAL PROCESSING LETTERS. Currently, he is an Associate Editor of IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS and a member of IEEE International Solid-State Circuits Conference International Program Committee serving in Imagers, MEMS, Medical, and Displays (IMMD) Subcommittee.

Antoine Adamantidis received pre- and postdoctoral education at the University of Liege, Liege, Belgium, and Stanford Medical School, Stanford, CA, USA. Currently, he is an Assistant Professor in the Department of Neurology, University of Bern, Bern, Switzerland, and holds a joined appointment in the Department of Clinical Research. He is the Codirector for the Zentrum for Experimentale Neurologie (ZEN labs) at the Inselspital in Bern. Also, he is an Adjunct Professor in the Department of Psychiatry at McGill University, Montreal, QC, Canada. His research interests aim at understanding the wiring diagram and the dynamics of sleep-wake circuits in the mammalian brain. His research objectives aim at investigating the wiring, firing dynamics, and plasticity of the neural circuits regulating brain states in normal and pathological states using in vitro and in vivo optogenetics—a technology that he and his colleagues pioneered at Stanford University—combined to genetics and electrophysiological methods. In his recent work, his laboratory identified a rapid-eye movement (REM) sleep circuit in the hypothalamus that control switch and duration of REM sleep states in mammals.

Dr. Adamantidis has received several awards, including the R. Broughton Young Investigator Award (Canadian Sleep Society), a Canadian Research Chair in Neural circuits and Optogenetics, a NIH Pathway to Independence (PI) Award-K99/R00 (USA), NARSAD, and Sleep Research Society Young Investigator Award (USA).