




Online Obstructive Sleep Apnea Detection on Medical Wearable Sensors

Grégoire Surrel , Amir Aminifar , Francisco Rincón, Srinivasan Murali, and David Atienza , *Fellow, IEEE*

Abstract—Obstructive Sleep Apnea (OSA) is one of the main under-diagnosed sleep disorder. It is an aggravating factor for several serious cardiovascular diseases, including stroke. There is, however, a lack of medical devices for long-term ambulatory monitoring of OSA since current systems are rather bulky, expensive, intrusive, and cannot be used for long-term monitoring in ambulatory settings. In this paper, we propose a wearable, accurate, and energy efficient system for monitoring obstructive sleep apnea on a long-term basis. As an embedded system for Internet of Things, it reduces the gap between home health-care and professional supervision. Our approach is based on monitoring the patient using a single-channel electrocardiogram signal. We develop an efficient time-domain analysis to meet the stringent resources constraints of embedded systems to compute the sleep apnea score. Our system, for a publicly available database (PhysioNet Apnea-ECG), has a classification accuracy of up to 88.2% for our new online and patient-specific analysis, which takes the distinct profile of each patient into account. While accurate, our approach is also energy efficient and can achieve a battery lifetime of 46 days for continuous screening of OSA.

Index Terms—Long-term monitoring, Obstructive Sleep Apnea (OSA), Online detection, Real-time classification, Wearable sensor.

I. INTRODUCTION AND STATE OF THE ART

OBSTRUCTIVE Sleep Apnea (OSA) is a common sleep disorder involving partial or complete obstruction of the upper airway. In the U.S. alone, Young *et al.* and Kapur *et al.* estimated that 3.8 million people between 30 and 60 years old are affected by this condition [1], [2]. Depending on the population lifestyle, the prevalence of OSA ranges from 3% to 24% according to Young *et al.* [3], with an estimated 5% worldwide by Kim *et al.* along with Lam *et al.* [4], [5]. This disorder is an aggravating factor for multiple health diseases, where Gaisl *et al.*, Peppard *et al.* and Yaggi *et al.* documented cardiovascular ones [6] (high blood pressure [7], stroke [8]). Schröner

Manuscript received November 16, 2017; revised February 24, 2018; accepted March 17, 2018. This work was partially supported in part by the Hasler Foundation under Grant 15048, in part by the ObeSense under Grant 20NA21 143081 RTD project, in part by the Swiss NSF, and in part by the Nano-Tera.ch with Swiss Confederation financing. This work was recommended by Associate Editor A. Humeau-Heurtier. (*Corresponding author: Grégoire Surrel.*)

G. Surrel, A. Aminifar, and D. Atienza are with the Embedded Systems Laboratory, EPFL, Lausanne 1015, Switzerland (e-mail: gregoire.surrel@epfl.ch; amir.aminifar@epfl.ch; david.atienza@epfl.ch).

F. Rincón and S. Murali are with the Smartcardia, Lausanne 1015, Switzerland (e-mail: francisco.rincon@smartcardia.com; srinivasan.murali@smartcardia.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TBCAS.2018.2824659

and O'Hara linked OSA with clinical depression [9] while Durmer *et al.* showed evidence of decreased memory and cognitive skills [10]. Because of the cardiac misbehaviors, people with OSA present a higher rate of sudden deaths, as proved by Gami *et al.* [11]. While this condition is treatable, Young *et al.* estimate that 90% of the subjects go undiagnosed [12]. Hence, there is a need for accessible obstructive sleep apnea screening.

Despite major progress, there is still a need to develop a non-intrusive solution for home OSA monitoring, for two reasons. First, there is low incentive for patients with low to moderate OSA to use external breathing equipment, such as Adaptive Servo Ventilation (ASV) [13]. Second, existing solutions are bulky, time-consuming, expensive and intrusive, as stated by Shouldice *et al.*, Marcos *et al.* as well as Koley and Dey [14]–[16]. Because the population-wide capacity of performing full OSA diagnosis does not match the recommended capacity from Flemons *et al.* [17], OSA testing or screening are only available in dedicated facilities for the most severe cases: the patient is required to go to a sleep center or hospital where his or her sleep will be monitored extensively for two non-consecutive nights. A full polysomnography (PSG) will be done to record electrocardiogram (ECG), electromyogram (EMG), electroencephalogram (EEG), eye movements, nasal airflow [18]. Altogether, it requires 22 electrodes plus a respiration mask. This is an intrusive setup that disturbs the patient's sleep quality. Moreover, the acquired data needs to be afterwards analyzed by a specialist. Given these constraints, Young *et al.* report that more than 80% of patients are reluctant to undergo a PSG [12]. Additionally, both the risks associated with an external respiratory help and the impracticality of OSA screening leaves most of the affected population without any kind of monitoring [12], hence the need for a simple yet efficient Internet of Things (IoT) solution usable for home screening with a possible doctor supervision.

In the context of non-intrusive OSA monitoring, previous studies have shown that it is possible to detect OSA based on single-lead ECG recordings, which has first been demonstrated in 1984 by Guilleminault *et al.* in [19]. The existing wearable devices for OSA detection, either commercially available or at the research state, from Bsoul *et al.*, Kelly *et al.*, Jarvis and Mitra, Raymond *et al.*, De Chazal *et al.*, McNames and Fraser, Stein and Domitovich, Mietus *et al.*, Shinal *et al.*, Drinnan *et al.*, Maier *et al.*, Schrader *et al.*, and Da Silva Pinho *et al.* [20]–[37], focus only on signal acquisition. However, the processing and identification of OSA events comes afterwards, as an additional offline phase, after downloading the data to a more powerful

platform. Nonetheless, due to the high rate of sudden death of people with OSA because of additional cardiac causes, as shown by Gami *et al.* [11], there is a clear need for personal real-time systems. Devices which integrate in the bed cannot currently access the problem of cardiac monitoring, and are not working well when two persons are in the same bed. Therefore, as emphasized by Fan *et al.* as well as Prathap *et al.* [38], [39], non-intrusive personal wearable systems need to be investigated.

In terms of OSA detection techniques and classification accuracy, the best one from a single-lead ECG recording reported in the literature is by McNames and Fraser, reaching 92.5% classification accuracy [28]. It is obtained by considering multiple features extracted from the frequency-domain as well as the ECG morphology. However, the main drawback of this solution is that the classification is manual and time-consuming. In fact, as a comparative figure, the accuracy of a manual classification from an expert using the full polysomnogram signal is, according to De Chazal *et al.*, 93% [40]. As opposed, among the fully automatic classification techniques, the approach proposed by De Chazal *et al.* [36] achieves the best results with an accuracy of 90.6%. When using exclusively time-domain features, non-linear statistics reach an accuracy of 85.6% along with 72.1% sensitivity and 91.2% specificity, as demonstrated by Maier *et al.* [33]. An "if-then" decision tree is considered by Fan *et al.* [38] with a reported accuracy of 93.2%, relying on features derived from the heart-beats. Nevertheless, it has the drawback of using an undefined subset of the database, so the results cannot be compared as-is to the previous numbers.

Finally, neural networks have also been considered for OSA classification by Da Silva Pinho *et al.* and show a performance of 82.1%, a sensitivity of 88.4% and a specificity of 72.3% [37], which is on par with the algorithms mentioned previously. Two groups, Da Silva Pinho *et al.*, and Pathinarupothi *et al.* [37], [41] have considered a similar approach using LSTM-RNN which report 82.1% up to 100% but they consider only a limited set of recordings which does not enable us to replicate the results using the full database.

However, among all the devices and techniques available, none is a high-accuracy wearable system with real-time and long-term screening while satisfying the need of cardiac monitoring. In [42], we proposed a low-power wearable system for real-time OSA screening along with cardiac monitoring. In this paper, we extend our previous work by improving both the classification accuracy and battery lifetime. In summary, we design in this work a wearable system for automatic, energy-efficient, non-intrusive, yet accurate OSA detection and screening using a single-channel ECG recording. Our main contributions are:

- We propose an online Obstructive Sleep Apnea detection technique compatible with cardiac-monitoring as developed by Sopic *et al.*, as well as De Giovanni *et al.* [43]–[45], with a time-complexity of $O(n)$, which is the theoretical lower bound. This is achieved by developing our own efficient outlier removal and through performing sleep apnea assessment in time-domain, which removes the need for computationally expensive frequency-domain analysis (see Section IV).

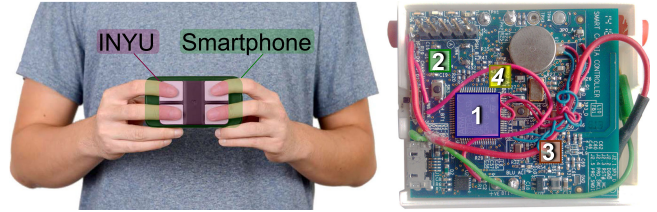


Fig. 1. INYU sensor and prototype. Front: 1: STM32L151RDT6 (ARM Cortex-M3 MCU, 384 KB Flash, 48 KB RAM), 2: MPU-6000 (6-axis I^2C motion sensor), 3: nRF8001 (Bluetooth low energy v4.0 radio). Back: 4: ADS1191 (Analog front-end for ECG applications).

- We design an autonomous and energy efficient sleep apnea screening system with a battery lifetime for continuous monitoring during 46 days, measured experimentally on an actual hardware platform (see Section VI-E).
- Our system, while energy efficient, is also comparable with the state-of-the-art in terms of performance, reaching 88.2% accuracy, 80.0% sensitivity and 93.9% specificity ($F1 = 84.7\%$) thanks to adopting a patient-specific perspective, which takes into account the distinct profile of each patient (see Section V-B).

The rest of this paper is organized as follows. First, we present the targeted hardware and software platform in Section II. Next, in Section III, we explain how features are generated and evaluated offline for optimizing the results in our online system. Then, in Section IV, we describe the implementation and optimization of our energy-efficient sleep apnea detection technique. In Section V, we define the setup used for testing our system, and then in Section VI, we evaluate our proposed system experimentally with respect to energy efficiency and classification accuracy, along with patient-specific configuration. Finally, in Section VII, we conclude that using our proposed patient-specific technique, it is possible to achieve high classification accuracy for OSA detection, while having a longer battery lifetime than the state-of-the-art systems.

II. SLEEP-APNEA MONITORING SYSTEM

In these sections, we first describe the wearable hardware platform and then present the software architecture.

A. Target Wearable Platform

We consider the SmartCardia INYU wearable sensor (Fig. 1) as our target device in this paper. INYU is an energy-efficient wearable device providing a single-lead ECG recording with a 24-bit ADC operating at a frequency from 250 Hz to 16 kHz. The ECG is measured using silver-chloride electrodes by impedance pneumography [46]. The device is equipped with the STM32L151RDT6 [47], an ultra-low power 32-bit microcontroller which can operate at a maximum frequency of 32 MHz. It features 48 kB of RAM and 384 kB of flash storage, and it is powered using 710 mAh battery.

Given the internal capabilities and connectivity possibilities, this device can work as a fully autonomous device for several days of continuous recording, uploading the recorded and pro-

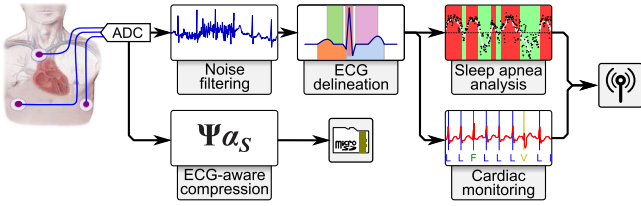


Fig. 2. Overview of the processing blocks integrated in the device used for the online OSA analysis in the proposed system [42]. The ECG is first filtered to remove the noise, then the fiducial points are extracted. Finally, signal is analyzed to detect OSA and cardiac pathologies. Raw ECG data is stored compressed on a memory for further offline analysis by an expert.

cessed data to a base station when one becomes available. For example, a Bluetooth Low-Energy compliant smartphone can be used for this purpose, which can afterwards display the data on-screen or upload it online to a remote medical service in order to be manually checked by a physician.

B. Software Structure

The overall software structure and flow of our proposed system is shown in Fig 2. The ECG is first acquired using medical electrodes connected to the chest of the patient. Then, an initial noise filtering is performed to remove artifacts caused, for example, by power lines, electrode parasitic motions, or baseline drift, as recommended by Webster and Huhta [48], [49]. Towards this, *erosion* and *dilation* morphological filters developed by Sun *et al.*, as well as Braojos *et al.* are used [50], [51], which can be implemented in an efficient way in energy-constrained wearable systems. After filtering the signal, the ECG delineation (extraction of fiducial points in the signal related to the physiological behavior) is done based on wavelet transforms following the method of Boichat *et al.* [52], relying on the fact that the different waves are made of different frequency components. The output of the previous step is then used to run two different automatic diagnostics. First, the OSA detection, which is the main focus of this paper and is discussed in Sections III and IV, but also a cardiac monitoring for additional evaluation of the patient. In parallel, the raw data is compressed using an algorithm from Mamaghanian *et al.* [53] and stored for further offline analysis.

III. OFFLINE FEATURE EXTRACTION AND OSA LEARNING PHASE

In this section, while keeping in mind the stringent energy-constraints of a wearable platform, we first identify relevant features in Subsection III-A and then, using these features, we describe how to train our classifier for OSA detection in Subsection III-B.

A. Features Extraction

Given that an autonomous device with online analysis has a longer battery life than when streaming the sampled data to a remote device such as a smartphone, provided that the embedded computations required for signal processing and classification

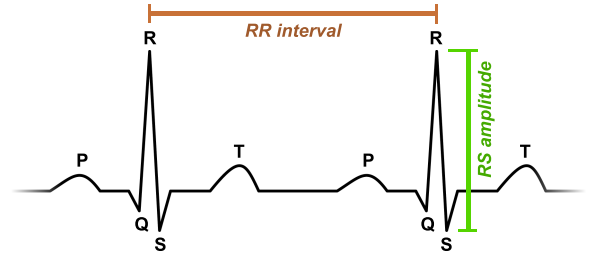


Fig. 3. ECG morphology with RR-interval and RS-amplitude labeled [59].

are lightweight enough, which has been proven by the works of Rincón *et al.*, Crepaldi *et al.*, and Basu *et al.* [54]–[57]. Thus, we need to carefully select the features we use for the OSA classification.

To select the most informative features, we evaluate the ones considered by De Chazal *et al.* [36]. This is because among the fully automatic systems relying solely on ECG, it reaches the highest accuracy (90%). In [36], 52 features are derived from the ECG morphology and an additional 36 features are derived from an EDR signal, i.e., a total of 88 features. In addition, we consider other features generated directly from the ECG, EDR, RR-intervals series and RS-amplitude series along with their respective spectrums. Besides statistical features (*min*, *max*, *mean*, *std*, *rmsd*, *sdnn*), we also consider auto-regressive process, i.e., $x_i = y_i - \frac{1}{4} \sum_{j=i-4}^{i-1} y_j$ as a previous study from De Chazal *et al.* [27] report an improved performance. Finally, we extract several features by integrating or deriving parts of the ECG signal. Therefore, a trade-off exists between the energy consumed for extracting the features and the classification performance. As our goal is to design an energy-efficient wearable system, it is essential to significantly reduce the number of features extracted and used.

To identify which features are the most relevant for OSA classification, we first generate all the features we consider from both the training and testing set. To select the features for the final system, we run a minute-by-minute classification, where the classification accuracy identifies how relevant each feature is for OSA detection. We apply forward feature selection as explained by Tang *et al.* [58] until the improvements are minimal: first, we select the one that gives the best accuracy when used alone. We then iterate, adding the next feature that, combined with the previously selected ones, provides the best classification accuracy.

As reported by Penzel *et al.* in a comparison of different algorithms for apnea detection from ECG recordings, the most common features are generated from the time series of heart beats, the ECG morphology, and from the ECG derived respiration (EDR) signal [60]. In particular, during an OSA event, there is a shift of the signal's energy towards low frequencies for two distinct time-series: the series of time intervals between two heart beats (RR-intervals in Fig. 3) and the series of R-peak amplitudes with respect to S-amplitudes (RS-amplitude in Fig. 3). This shift of signal's energy is illustrated in Fig. 4 with the spectrogram of the RR-intervals series. It shows the minutely frequency-spectrum of the RR-intervals for a complete

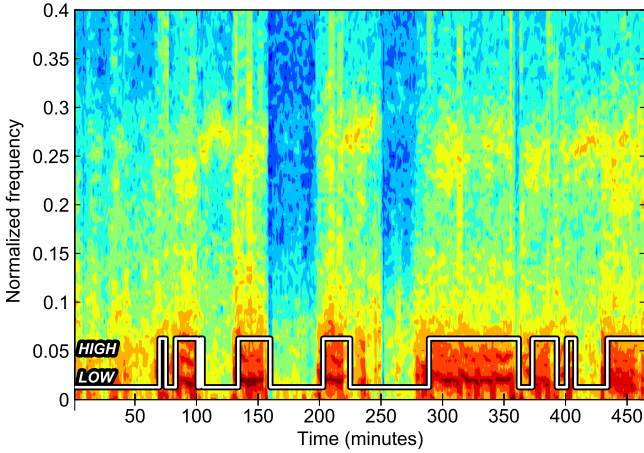


Fig. 4. Spectrogram of the log-power of the RR-intervals series from the recording x_{32} (see Section V-A) with the labeled OSA logic signal on a same time axis [42].

overnight recording. Frequencies of the signal are displayed on the y-axis versus the time on the x-axis. As for the background, the dark red color indicates high energy for that frequency and minute, whereas the blue is linked to low energy. The manually-labeled OSA ground truth by the medical expert is shown below. In these annotations, the *HIGH* value is linked to OSA events whereas normal breathing is captured by the *LOW* value. As this figure shows, there is a clear correlation between the OSA events and the signal's energy in low frequencies. A very similar trend is observed for the RS-amplitude series. As a result, we consider both RR-intervals spectrum and the RS-amplitude spectrum as both are correlated with OSA.

Computing signal's spectrum is, however, energy-hungry. We, therefore, isolate the most relevant frequency-band from each spectrum (RR-interval and RS-amplitude). Thus, we run a parameter sweep to optimize the frequency-band bounds of the signal correlated with OSA events, with respect to classification accuracy. From an exhaustive exploration of all possible frequency bands, we obtain a 2D-map of accuracy (see Fig. 5). It represents the OSA classification accuracy obtained for each pair of low and high bounds. The apnea frequency-band bounds are on the axes and the classification accuracy is linked to the graph color: the brighter the color, the higher the accuracy. This figure, which is computed for RR-intervals, shows a clear frequency interval regarding the lower bound of the band, ranging from 0.010 to 0.020, while the upper bound has more tolerance, from 0.045 to 0.075, in terms of normalized frequency. This whole area provides the highest classification results, reaching 76% when using the raw RR-intervals without filtering. Therefore, we have a single feature from the spectrum of RR-intervals for classification, which can be computed in an energy-efficient way (cf. Section IV-B). Similarly, the best frequency-band bounds are found for the RS-amplitude time-series, which leads us to an accuracy of 76%.

B. Features Combination and Learning Phase

Based on the discussion in the previous section, we only use the relative energy in a specific frequency band for both RR-intervals and RS-amplitudes time-series, as they are two features

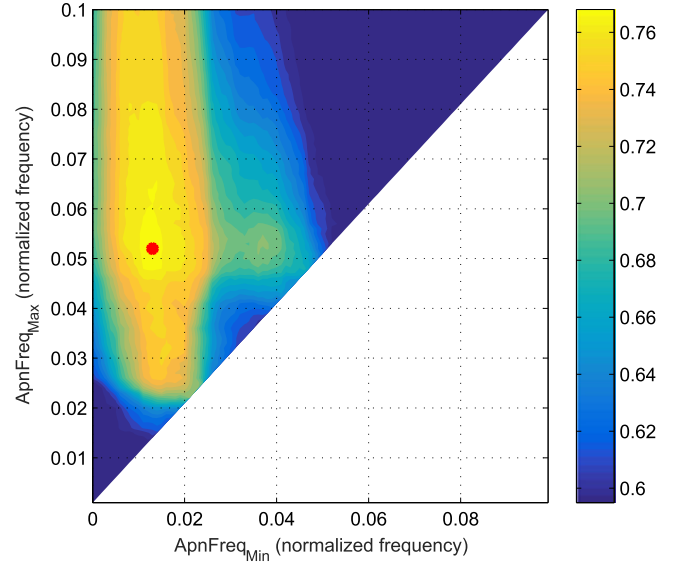


Fig. 5. Classification accuracy (normalized between 0 and 1) for OSA when varying frequency-band bounds considering the RR-intervals time series. The circular dot is placed at the position of the best normalized frequency band.

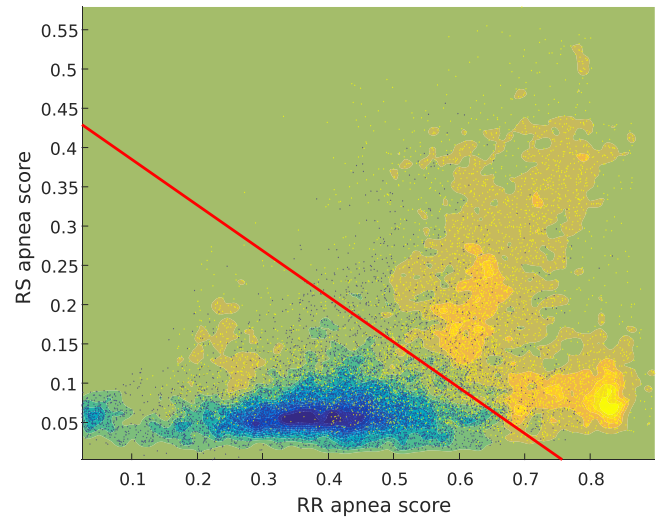


Fig. 6. Distribution of the *apnea* (yellow) and *non-apnea* (blue) events from the training set of recordings for the two features generated from the series of RR-intervals and RS-amplitudes. The two classes are separated based on linear SVM classifier.

not correlated to each other. This is illustrated in Fig. 6 where each sample data is positioned in the RR-RS apnea score plot. The sample data for OSA events are in yellow and normal sleep are in blue. The red line (linear combination of RR and RS apnea scores) separates the two classes by maximizing the margin.

In the literature, several classifiers were considered and the classification accuracy obtained is in the range between 80% and 90%, relying from the numbers reported by Xie and Minn [61]. Additionally, Xie and Minn show in [61] that Random Forest [62] has the highest performance in terms of classification accuracy. In our setup, Random Forest reaches a classification accuracy of 88.1% at most. On the other hand, when using a Support Vector Machine (SVM) classifier as defined by Cortes and

Vapnik [63], we reach 90.1% classification accuracy, whether we use a linear, Gaussian or polynomial kernel. Therefore, in terms of classification accuracy, the results obtained using the SVM classifier are up to 2% better than Random Forest's results.

In terms of computational complexity, linear SVM can be implemented even with limited computational resources. As we design an energy-constrained system, we use linear SVM in our OSA detection technique, because of its high classification performance and high computational efficiency at runtime (see Section IV-D).

IV. ONLINE SLEEP-APNEA DETECTION TECHNIQUE

In this section, we propose an energy efficient Obstructive Sleep Apnea detection technique that can directly run on an energy-constrained wearable device. Previous work and devices from Jarvis and Mitra, Raymond *et al.*, De Chazal *et al.*, McNames and Fraser, Stein and Domitovich, Mietus *et al.*, Shinar *et al.*, Drinnan *et al.*, Maier *et al.*, Schrader *et al.*, Ng *et al.*, Da Silva Pinho *et al.*, Kelly *et al.*, and Bsoul *et al.* [23]–[37] focused on offline analysis on a computer or cloud system. Conversely, we aim to provide an online ECG analysis running on a wearable system. Therefore, our main goal here is to lower the energy consumption of the OSA detection, while maintaining high classification accuracy on the wearable device.

The overall flow of our online sleep apnea analysis, after the ECG noise filtering and ECG delineation, is shown in Fig. 7. Our OSA screening method allows different lengths of interval but in particular in this paper, we use a 60 seconds analysis as the database we use provides only a minute-by-minute labeling. From the ECG, we generate RR-intervals and RS-amplitudes time-series, as they are our most relevant features (see Section III). First, we use our own enhanced Thompson-Tau filter (Section IV-A) to remove the outliers from erroneous beats caused mostly by motion artifacts and muscle noise, inherent to ambulant systems. Then, we compute the power in the two apnea frequency bands of the spectrums in order to obtain the RR apnea score as well as the RS apnea score (Section IV-B). Afterwards, we apply a moving average filter to smooth the variability of both apnea scores (Section IV-C). Finally, according to the SVM classifier trained in Section III-B, if the linear combination of the smoothed apnea scores is greater than a threshold, we label the corresponding minute as apnea.

A. Low-Complexity Outlier Removal

As proved by Clifford and Tarassenko heart-beat outliers have strong negative effects on frequency analysis [64]. Consequently, detecting and removing these outliers is critical. Therefore, we apply our own low-complexity version of the Thompson-Tau filter on the series of RR-intervals before apnea scoring. The original Thompson-Tau outlier removal algorithm is provided by Rienzner [65] and its average time complexity is $O(n \log(n))$, because of sorting the entire input array. However, we propose faster outlier removal by reducing the average time complexity down to $O(n)$ (see Algorithm 1). Our algorithm first starts by computing the initial mean and standard-deviation

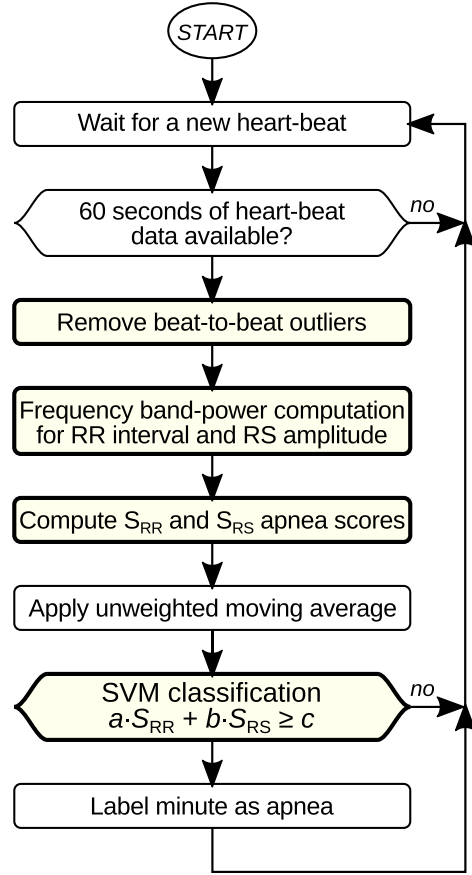


Fig. 7. Flowchart of the proposed online OSA screening technique. It is applied when considering both the RR-intervals and the RS-amplitudes of the beats. Compared to [42], the highlighted blocks feature the use of our optimized outliers removal, updated apnea-score computation using a time-domain filtering and power estimation, as well as the additional RS-amplitude feature extraction along with the inclusion of the SVM classifier.

of the series (Lines 3–4). In Lines 5–6, it finds the k -th largest and the k -th smallest values using the QuickSelect algorithm documented by Hoare [66], which only sorts the beginning and ending of the input array. In the subsequent Lines 8 and 13, we test if the smallest or largest value is conformant to the Student's-t distribution [67] of the input. If not (Lines 9–12 and 14–17), then we move the *start* or *end* indices of the series to exclude the new outlier and update both the mean and standard-deviation according to the equations from Welford [68]. We repeat this process until both the smallest and largest values are conforming to the distribution (Lines 7–19).

The average complexity of Algorithm 1 is $O(n)$. In the impossible case where all values are outliers, the complexity is the same as in the original implementation. However, only a few values are outliers. Indeed, Fig. 8 shows the classification accuracy while changing the filter's tolerance for outliers detection. As this figure shows, there is a 5% increase of the classification accuracy if a very tolerant filter is used, thus discarding the few biggest outliers. It means that removing a restricted number of the most non-conforming samples brings significant improvements. We can, therefore, abort sorting early, thus providing significant energy savings.

Algorithm 1: Low-Complexity Outlier Removal.

```

1: function FASTOUTLIERREMOVAL( $\mathbf{x}, \tau, k$ )
2:    $N \leftarrow \text{LENGTH}(\mathbf{x});$    $start \leftarrow 1;$    $end \leftarrow N;$ 
3:    $\mu \leftarrow \text{COMPUTEMEAN}(\mathbf{x});$ 
4:    $\sigma \leftarrow \text{COMPUTESTD}(\mathbf{x});$ 
5:    $low \leftarrow \text{QUICKSELECT}(\mathbf{x}, k);$             $\triangleright$  Sorts the
beginning of the array
6:    $high \leftarrow \text{QUICKSELECT}(\mathbf{x}, N - k);$     $\triangleright$  Sorts
the ending of the array
7:   while  $\mathbf{x}[start] < low$  or  $\mathbf{x}[end] > high$  do
8:     if  $low > \tau \cdot \sigma$  then
9:        $\mu \leftarrow \text{UPDATEMEAN}(\mathbf{x}[start], N, \mu);$ 
10:       $\sigma \leftarrow \text{UPDATESTD}(\mathbf{x}[start], N, \sigma, \mu);$ 
11:       $start \leftarrow start + 1;$ 
12:       $N \leftarrow N - 1;$ 
13:     else if  $high < \tau \cdot \sigma$  then
14:        $\mu \leftarrow \text{UPDATEMEAN}(\mathbf{x}[end], N, \mu);$ 
15:        $\sigma \leftarrow \text{UPDATESTD}(\mathbf{x}[end], N, \sigma, \mu);$ 
16:        $end \leftarrow end - 1;$ 
17:        $N \leftarrow N - 1;$ 
18:     end if
19:   end while
20: end function

```

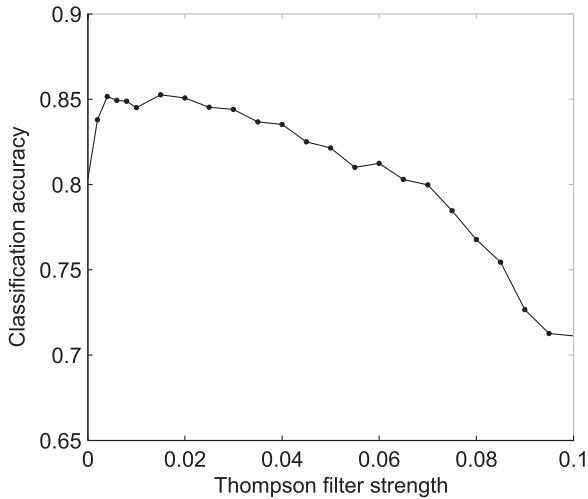


Fig. 8. Evolution of the accuracy depending on the strength of the Thompson filter. Even a very tolerant (weak) removal significantly improves the classification accuracy, which means that only removing the few biggest outliers is desirable [42].

B. Apnea Scoring

In this subsection, we define two apnea scores and discuss our time-domain analysis to compute both of them. They are the features used by the linear SVM for classifying OSA events.

Definition: We define the apnea score S as the relative energy in the apnea band $E_{\text{apnea-band}}$ compared to the total signal energy E_{total} :

$$S = \frac{E_{\text{apnea-band}}}{E_{\text{total}}}. \quad (1)$$

This enables us to consider the apnea score computed on the series of RR-intervals S_{RR} with the corresponding apnea band $E_{\text{RR-apnea-band}}$ as well as the apnea-score computed on the series of RS-amplitudes S_{RS} with the corresponding apnea band $E_{\text{RS-apnea-band}}$. Note that the apnea score is bounded between zero and one.

To compute the energy in given frequency bands, a Lomb normalized periodogram is typically used, using the method of Karakonstantis *et al.* [69], but the drawback of this approach is the complexity associated with the generation of the whole frequency power spectrum at run time. Indeed, the frequency transform yields as many features as discrete frequencies, which involves energy-hungry computations. Contrarily, in our case, to compute an apnea score S , we only need to compute the energy of a signal in the corresponding apnea band $E_{\text{apnea-band}}$, as well as the total energy of the signal E_{total} . Therefore, it is possible to substantially lower the memory and CPU usage and save energy without sacrificing the classification accuracy.

To compute the the total energy of the signal E_{total} , we rely on Parseval's Theorem [70], and compute the signal's energy E_{total} using lightweight time-domain signal processing, instead of time-domain to frequency-domain transforms:

$$\sum_{n=0}^{N-1} |\mathbf{x}(n)|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |\mathbf{X}(k)|^2, \quad (2)$$

where $\mathbf{X}(\cdot)$ is the Discrete Fourier Transform (DFT) of signal $\mathbf{x}(\cdot)$.

For the energy of the signal in the apnea band $E_{\text{apnea-band}}$ we first design a time-domain band-pass filter to remove the frequencies outside the apnea band. A first-order digital Butterworth band-pass filter [71] has been designed to compute the coefficients a_i and b_i of the transfer function:

$$H(z) = \frac{\sum_{i=0}^m b_i z^{-i}}{\sum_{i=0}^k a_i z^{-i}}. \quad (3)$$

Using these coefficients, we obtain an energy-efficient time-domain filter. The time-domain Infinite Impulse Response (IIR) digital filter is given as follows:

$$a_0 \mathbf{y}(n) = \sum_{i=0}^m b_i \mathbf{x}(n-i) - \sum_{i=1}^k a_i \mathbf{y}(n-i). \quad (4)$$

Having filtered the frequencies outside the apnea band, we can again use Parseval's Theorem to obtain the apnea band energy $E_{\text{apnea-band}}$. Once this is done, apnea score is computed according to the definition given in Equation 1.

Considering the time-domain energy computation, along with the time-domain band-pass signal filtering, the apnea-scoring is more efficient from the energy-consumption point of view because of the reduced algorithmic complexity, compared to computing our apnea-score using a frequency domain transform.

C. Apnea-score Low-Pass Filtering

The reference minute-by-minute sleep-apnea labels reveal that OSA is a signal that changes at low frequencies. In fact, it

is unlikely to have a single minute containing an apnea event in a long period of non-apnea sleep and vice-versa. Therefore, we consider the evolution of the apnea score over several minutes. Hence, we use a simple unweighted moving average to lower the raw apnea score variability, as follows:

$$\mathbf{x}(i) = \frac{1}{2m+1} \sum_{j=-m}^m \mathbf{x}(i+j). \quad (5)$$

We optimize the moving average window length by maximizing the OSA classification on the training set when changing the window length from zero minutes up to 31 minutes. The classification accuracy based on the RR apnea score increases by 3–4%, if we consider a window length of 13 ± 4 minutes. Similarly, in the case of RS apnea score, the best improvement in accuracy is achieved for a five minutes window.

D. Online Classification

Once the signal is filtered, the features generated and filtered, the final step is to actually label the minute as apnea or not. We classify each minute of signal using a linear Support Vector Machine (SVM) using the parameters from the offline training done in Section III. It is especially energy-efficient as it only requires computing the following condition:

$$a \cdot S_{RR} + b \cdot S_{RS} \geq c,$$

with the parameters a , b and c computed during an offline training (cf. Section III-B).

V. EXPERIMENTAL SETUP

In this section, we define the database used as well as the setup for testing the performance of our system.

A. Apnea-ECG Benchmark Recordings

The recordings used for training and testing are publicly available on Physionet as the apnea-ecg database [72]. They were made available by Penzel *et al.* for stimulating research about non-intrusive OSA detection. We use this database to be able to compare our proposed system against prior studies, as we report our results using the same experimental methodology and metric based on the same signals. The 70 single-lead ECG recordings were sampled at a frequency of 100 Hz and manually labeled minute-by-minute by an expert for sleep apnea and hypopnea events, without distinction between both. The heart-beat timestamps used to retrieve the RR-intervals are provided by Physionet using an automatic delineation.

The recordings come from a set of 32 subjects, namely, healthy and with Obstructive Sleep Apnea. From those subjects, four subjects contributed to four recordings each, two subjects contributed to three recordings each, 22 subjects contributed to two recordings each and four patients contributed to a single recording. Then, the dataset is divided in two groups of 35 recordings, one for training and one for testing. In each group, the number of apnea events represent around 38% of the data. The total number of recorded minutes is 34313, and we include

17045 of them in the training set and the remaining 17268 in the testing set.

The duration of the recordings ranges from 6 h 41 min to 9 h 38 min, with an average duration of 8 h 12 min. The normal breathing time varies between 11 and 535 minutes, whereas for the problematic breathing, it ranges from 0 to 534 minutes. Overall, 62% of the minutes in the database are labeled as apnea. This means that a system classifying everything as apnea would only reach 62% accuracy. The amount of breathing-disordered minutes is used to classify the patients in three different groups: the apnea group A was defined as having 100 or more minutes of Obstructive Sleep Apnea and the control group C showed less than 5 minutes of disordered breathing. The remaining cases belong to group B, classified as borderline, i.e., with between 5 and 99 minutes with apnea during the recording.

B. OSA Classification Performance

To have a performance comparison with prior works, we use the overall classification accuracy when working on the testing set recordings. Our OSA monitoring system performs a minute-by-minute analysis, assigning either the *non-apnea minute* label or the *apnea minute* one. We also provide the specificity and sensitivity to fully characterize our system. These indicators are defined as follows:

Sensitivity (or True Positive Rate):

$$TPR = \frac{TP}{TP + FN} = \frac{TP}{RP}, \quad (6)$$

Specificity (or True Negative Rate):

$$TNR = \frac{TN}{FP + TN} = \frac{TN}{RN}, \quad (7)$$

Accuracy:

$$Acc = \frac{TP + TN}{RP + RN}, \quad (8)$$

where TP is the number of true positives (correctly classified minute as apnea), TN is the number of true negatives (correctly classified minute as non-apnea), FP is the number of false positives (misclassified minute as apnea), FN is the number of false negatives (misclassified minute as non-apnea), and RP and RN are, respectively, the number of real positives (apnea minutes from the ground truth) and real negatives to classify (non-apnea minutes from the ground truth).

VI. EXPERIMENTAL RESULTS

In this section, we evaluate our approach in terms of classification performance: The following four subsections report the classification accuracy under different assumptions.

First, Subsection VI-A reports the performance using exactly the the same setup than the Physionet Challenge. This enables an easy comparison with other work using the same database. The next three subsections rely on the patient-specific grouping of recordings. To evaluate our patient-specific approach, the recordings are grouped by patient using the metadata provided along the database, using the reported age, sex, height

TABLE I
GROUPING OF OBSTRUCTIVE SLEEP APNEA RECORDINGS ON
A PER-PATIENT BASIS

Patient	Recordings	Patient	Recordings
01	a01, a14	17	b05, x11
02	a02, x14	18	c01, x35
03	a03, x19	19	c02, c09
04	a04, a12	20	c03, x04
05	a05, a10, a20, x07	21	c04, x29
06	a06, x15	22	c05, x33
07	a07, a16, x01, x30	23	c06
08	a08, a13, x20	24	c07, x34
09	a09, a18	25	c10, x18
10	a11	26	x02
11	a15, x27, x28	27	x06, x24
12	a17, x12	28	x09, x23
13	a19, x05, x08, x25	29	x10
14	b01, x03	30	x13, x26
15	b02, b03, x16, x21	31	x17, x22
16	b04, c08	32	x31, x32

and weight, without ambiguity. The results of this grouping are shown in Table I, sorted in the same order as the original files. As some patients (patients 10, 23, 26 and 29) contributed to a single recording, they could not be used for the patient-specific study and, therefore, have been excluded. This is because we require one recording for training, and at least another recording for accessing the classification performance. Our results from the patient-specific approach are reported in Table II.

A. *Physionet Challenge Classification Accuracy*

In this work, we consider the exact same setup defined by Moody *et al.* for the Physionet Challenge [73]. After the optimization and porting to embedded C, the accuracy we reach is 85.7% while the sensitivity is 81.4% and specificity is 88.4% ($F1 = 81.3\%$). If we only consider the RR-intervals series, the accuracy is 82.2%, with 73.3% sensitivity and 87.6% specificity ($F1 = 75.7\%$). These results are better by 0.1% of the best time-domain classifiers, by Maier *et al.* [33], and 3 to 5% lower to the absolute best algorithms, from Da Silva Pinho *et al.*, and De Chazal *et al.* [36], [37], with the added benefit of online energy-efficient analysis on a wearable device.

B. *Patient-ideal (P.I.) Classification Accuracy*

In order to quantify how much improvement can be expected from the patient-specific approach, for each patient, training and testing is done on the same recordings. The overall classification accuracy achievable with our technique in this ideal case, where we have full knowledge of the recordings, is 91.3% (specificity = 92.7%, sensitivity = 89.2%). This is the best achievable performance in terms of classification accuracy with our OSA detection technique.

C. *Patient-agnostic (P.A.) Classification Accuracy*

To compare with a patient-specific approach, we group the recordings in Table I by patients. We use the first recording of each patient for training and the remaining recordings for

testing. We leave out four patients as they contributed to only a single recording. Using this approach, we reach a classification accuracy of OSA events of 84.5% (specificity = 88.1%, sensitivity = 79.2%).

D. *Patient-specific (P.S.) Classification Accuracy*

Even though the overall classification accuracy of obstructive sleep-apnea events is high (more than 85%), few patients present significantly lower classification results (below 70%). After grouping the recordings per patient, we observe that they have a distinct apnea profile for the energy distribution in the RR and RS apnea bands. We, therefore, propose a patient-specific SVM classifier training in our obstructive sleep apnea screening system.

In the patient-agnostic situation, we use 35 recordings coming from different patients for training, as per the Physionet Challenge rules, to have comparable results with the state of the art. In the patient-specific situation, we use the first overnight recording from the patient to train the SVM classifier.

With an exhaustive analysis across all the patients, we observe that, on the one hand, this patient-specific approach is overall 4.4% better than the patient-agnostic setup (specificity = 93.9%, sensitivity = 80.0%). On the other hand, our patient-specific OSA detection technique is 3% less accurate than the patient-ideal approach. Even though in a very few cases the performance is slightly degraded, this tuning brings significant improvements for the majority of patients with a very low accuracy in a patient-agnostic setting (see Table II and Fig. 9). This observation opens the possibility of a long-term patient screening for a wider range of patients.

On the other hand, let us consider, as an example, Patient 15 who has contributed to four recordings. Without any patient-specific training, all four recordings receive a poor classification results, as the initial patient-agnostic accuracy is 69.8%. However, when training using one recording from the patient, the classification accuracy increases to 88.2% (we avoid data-dredging by excluding the recording used training). These results are similar when using any of the recordings for training, and testing against the other recordings.

Let us now consider Patient 1. The classification accuracy is poor compared to the majority of patients (below 70% for both the patient-specific and patient-agnostic cases). To find the best performance achievable in theory, we both train and test our technique on the recordings from the patient. This ideal classification can achieve optimal accuracy as it is based on data-dredging. The ideal accuracy, under our assumptions, is 76.0%, which is comparable with the classification accuracy obtained by our patient-specific training (75.4%).

To give more insight about the performance of our system, we plot the number of used features against the classification accuracy. As only few publications report the number of used features, many publications are not in the chart. Apart from our system, the best performing one with the lowest number of features is by De Chazal *et al.* [36], with 87.7% accuracy using 9.7 features (non-integer value because it is the average number of features used in multiple data splits from a cross-validation

TABLE II
CLASSIFICATION ACCURACY AND RELATIVE IMPROVEMENTS REPORTED (IN %) ON A PATIENT-SPECIFIC BASIS

Patient	1	2	3	4	5	6	7	8	9	11	12	13	14	15	16	17	18	19	20	21	22	24	25	27	28	30	31	32	Avg.	Std.
<i>P.A.</i>	64	54	90	72	74	66	80	83	86	92	80	85	95	70	98	98	100	100	100	98	99	98	99	100	98	69	100	87	84	12
<i>P.S.-P.A.</i>	17	67	1	-2	6	11	-5	-1	5	-7	8	5	2	30	2	-0	0	0	0	2	1	1	0	0	-3	1	0	11	4	13
<i>P.A.</i> <i>P.S.</i>	75	90	91	70	78	73	76	82	90	85	86	89	97	91	100	97	100	100	100	100	99	99	100	100	95	70	100	96	88	10
<i>P.I.-P.S.</i>	1	0	1	32	4	14	5	4	-0	9	8	4	0	0	0	0	0	0	0	0	0	0	0	0	3	20	0	-0	3	7
<i>P.S.</i> <i>P.I.</i>	76	90	92	93	81	84	79	86	90	93	92	93	97	91	100	97	100	100	100	100	99	99	100	100	98	83	100	96	91	7

This table shows the change of the accuracy when changing from a patient-agnostic (*P.A.*) classifier to a patient-specific (*P.S.*) one and then to a ideal patient-specific (*P.I.*) classifier.

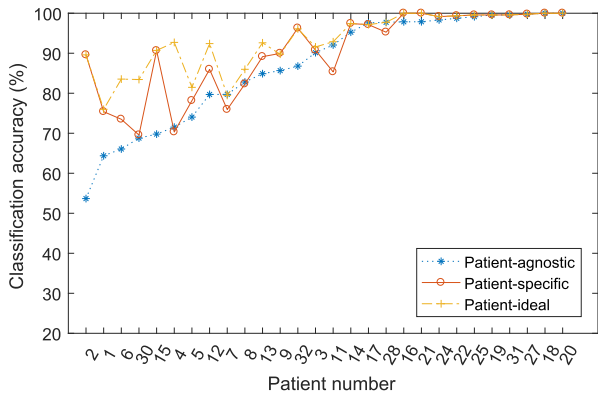


Fig. 9. Evolution of the OSA classification accuracy in the context of patient-agnostic, patient-specific and patient-ideal classification. The patients have been sorted according to the patient-agnostic classification accuracy to better show the classification accuracy changes with the patient-specific and patient-ideal cases.

process). That is to say, it has 3% better accuracy at the cost of more than four times as many features.

E. Energy Consumption Characterization

Targeting an autonomous hardware platform INYU, in addition to the classification performance, it is important to take energy consumption into consideration as a design goal. Therefore, in this section, we evaluate the energy efficiency of our system experimentally. We use the commercially available Gecko EFM32 development board (with the same ARM Cortex-M3 core as in the INYU) with the provided Simplicity Studio software as a full energy profiler is integrated.

The detailed energy analysis of the OSA detection algorithm (Fig. 11) has been performed on a recorded set of data, spanning over 12.3 hours, with an average heart-rate of 87 beats per minute. The total active time is 38.78 seconds, which represents a duty cycle of 0.085%. The average current drawn by the microcontroller while active is 10.5 mA.

In the case where the device is only used for OSA detection, the energy consumption results are reported in Table III. The ECG measurement is active 100% of the time. Concerning the software, two main parts are required: the ECG delineation to detect the heart-beat and the OSA detection. In both cases, the microcontroller is in its active state, drawing 10.5 mA. When idle, the CPU is in an energy saving mode, drawing 0.018 mA. By performing an analysis of the different individual consumptions, we get an average current consumption of 0.636 mA.

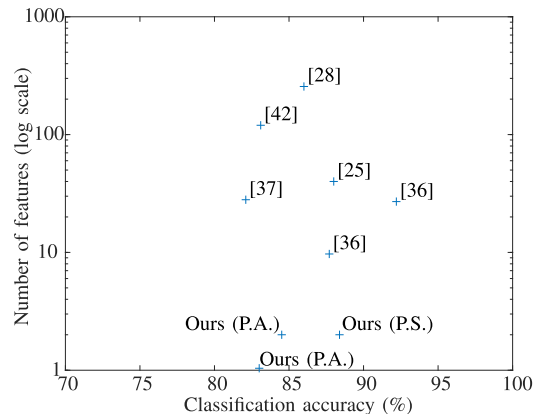


Fig. 10. Comparison between the the number of features used against the classification accuracy among the published papers.

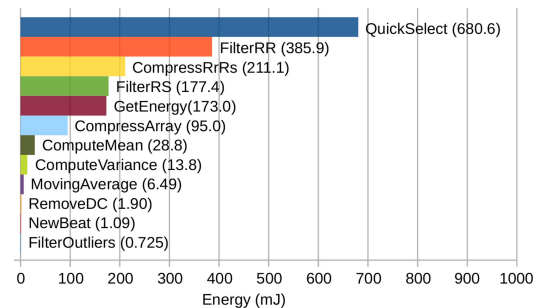


Fig. 11. Overview of the energy consumption of the OSA detection algorithm. Each bar represents the energy consumed by a single function in the C code running on the energy-profiling board.

TABLE III
CURRENT USED FOR OSA DETECTION ON THE TARGET DEVICE

Operation	Current (mA)	Duty cycle (%)	Avg. current (mA)
ADS1191 [74]	0.427	100	0.427
MPU-6000 [75]	0.005	100	0.005
Signal acquisition subsystem			0.432
ECG delineation	10.5	1.667	0.175
Apnea processing	10.5	0.085	0.009
Idle time	0.018	98.25	0.018
STM32 [76] data processing subsystem			0.202
nRF8001 [77]	11	0.0007	0.008
Wireless subsystem			0.008
Total			0.637

The currents drawn are based both on the manufacturer's datasheets and are confirmed experimentally with measurements. As the duty cycle is data dependent, we determine an average active time for various ECG windows extracted from the Physionet recordings, thus reflecting the variability observed for real signals.

As the battery is rated at 710 mAh, the total lifetime is approximately 1115 hours (46.5 days). Thus, the obstructive sleep apnea detection technique in this paper is one order of magnitude more energy-efficient than our previous work [42], while having better classification results whether we consider a patient-agnostic or patient-specific approach. This is mainly due to our fast outliers removal and time-domain apnea score computation.

VII. CONCLUSION

Obstructive sleep apnea is an aggravating factor for different health conditions, including cardiovascular diseases. Despite the high rate of obstructive sleep apnea, only a small fraction of the population is diagnosed and monitored. Therefore, in this paper, we designed an online ultra-low power wearable obstructive sleep apnea monitoring system. The performance and energy efficiency of our system are evaluated experimentally, in a patient-specific setting. Our system has a classification accuracy of 88.2%, for a minute-by-minute classification, with a battery lifetime of 46.8 days. Thanks to its Bluetooth link, this wearable sensor can upload its analysis to an online web-service for a continuous monitoring, tracking the evolution of the disease.

As for future work, we can envision two main directions. First, our proposed wearable system can be expanded to provide a better sleep diagnosis device, with the practical integration of the cardiac monitoring, using the cardiac analysis results to refine the OSA analysis. An additional promising evolution is the automatic disabling of the ECG sampling when a noisy section is detected, which can be performed by exploring as a basis the strategy proposed by Orphanidou *et al.* [78]. Second, personal activity trackers (fitness trackers and smartwatches) are getting widespread, and the majority of them feature a pulse oximeter, yielding a photoplethmogram (PPG) and therefore an indication of the heart-rate. Hence, it would be interesting to evaluate the performance of our OSA detection algorithm relying only on the RR-intervals measured with the PPG, thus using devices already available on the mass market.

REFERENCES

- [1] T. Young, M. Palta, J. Dempsey, J. Skatrud, S. Weber, and S. Badr, "The occurrence of sleep-disordered breathing among middle-aged adults," *New England J. Med.*, vol. 328, no. 17, pp. 1230–1235, 1993.
- [2] V. Kapur, K. P. Strohl, S. Redline, C. Iber, G. O'connor, and J. Nieto, "Underdiagnosis of sleep apnea syndrome in us communities," *Sleep Breathing*, vol. 6, no. 2, pp. 49–54, 2002.
- [3] T. Young, P. E. Peppard, and D. J. Gottlieb, "Epidemiology of obstructive sleep apnea: A population health perspective," *Amer. J. Respiratory Critical Care Med.*, vol. 165, no. 9, pp. 1217–1239, 2002.
- [4] J. Kim *et al.*, "Prevalence of sleep-disordered breathing in middle-aged korean men and women," *Amer. J. Respiratory Critical Care Med.*, vol. 170, no. 10, pp. 1108–1113, 2004.
- [5] J. C. Lam and M. S. Ip, "An update on obstructive sleep apnea and the metabolic syndrome," *Current Opinion Pulmonary Med.*, vol. 13, no. 6, pp. 484–489, 2007.
- [6] T. Gaisl, D. J. Bratton, and M. Kohler, "The impact of obstructive sleep apnoea on the aorta," *Eur Respiratory J*, vol. 46, no. 2, pp. 532–544, Jun. 2015. [Online]. Available: <http://dx.doi.org/10.1183/09031936.00029315>
- [7] P. E. Peppard, T. Young, M. Palta, and J. Skatrud, "Prospective study of the association between sleep-disordered breathing and hypertension," *New England J. Med.*, vol. 342, no. 19, pp. 1378–1384, May 2000. [Online]. Available: <http://dx.doi.org/10.1056/nejm200005113421901>
- [8] H. K. Yaggi, J. Concato, W. N. Kernan, J. H. Lichtman, L. M. Brass, and V. Mohsenin, "Obstructive sleep apnea as a risk factor for stroke and death," *New England J. Med.*, vol. 353, no. 19, pp. 2034–2041, Nov. 2005. [Online]. Available: <http://dx.doi.org/10.1056/nejmoa043104>
- [9] C. Schröder, and R. O'Hara, "Depression and obstructive sleep apnea (OSA)," *Ann. Gen. Psychiatry*, vol. 4, pp. 343–352, Jun. 2005.
- [10] J. S. Durmer and D. F. Dinges, "Neurocognitive consequences of sleep deprivation," in *Seminars in Neurology*, vol. 25, New York, NY, USA: Thieme Medical Publishers, Inc., 2005, pp. 117–129.
- [11] A. S. Gami, D. E. Howard, E. J. Olson, and V. K. Somers, "Day-night pattern of sudden death in obstructive sleep apnea," *New England J. Med.*, vol. 352, no. 12, pp. 1206–1214, 2005.
- [12] T. Young, L. Evans, L. Finn, and M. Palta, "Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women," *Sleep*, vol. 20, no. 9, pp. 705–706, 1997.
- [13] "Adaptive servo ventilation," 2015. [Online]. Available: <https://www.sleepassociation.org/adaptive-servo-ventilation/>
- [14] R. B. Shouldice, L. M. O'Brien, C. O'Brien, P. de Chazal, D. Gozal, and C. Heneghan, "Detection of obstructive sleep apnea in pediatric subjects using surface lead electrocardiogram features," *Sleep*, vol. 27, no. 4, pp. 784–792, 2004.
- [15] J. V. Marcos, R. Hornero, I. T. Nabney, D. Álvarez, and F. Del Campo, "Analysis of nocturnal oxygen saturation recordings using kernel entropy to assist in sleep apnea-hypopnea diagnosis," in *Proc. IEEE Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2011, pp. 1745–1748.
- [16] B. L. Koley and D. Dey, "Real-time adaptive apnea and hypopnea event detection methodology for portable sleep apnea monitoring devices," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 12, pp. 3354–3363, Dec. 2013.
- [17] W. W. Flemons, N. J. Douglas, S. T. Kuna, D. O. Rodenstein, and J. Wheatley, "Access to diagnosis and treatment of patients with suspected sleep apnea," *Amer. J. Respiratory Critical Care Med.*, vol. 169, no. 6, pp. 668–672, 2004.
- [18] A. A. of Sleep Medicine, and C. Iber. (2007). *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. Darien, IL, USA: American Academy of Sleep Medicine.
- [19] C. Guilleminault, R. Riley, and N. Powell, "Obstructive sleep apnea and abnormal cephalometric measurements: Implications for treatment," *Chest*, vol. 86, no. 5, pp. 793–794, 1984.
- [20] N. Oliver and F. Flores-Mangas, "HealthGear: A real-time wearable system for monitoring and analyzing physiological signals," in *Proc. Int. Workshop Wearable Implantable Body Sensor Netw.*, 2006, pp. 4–64.
- [21] "SleepImage," 2016. [Online]. Available: <http://www.sleepimage.com/>
- [22] "Sleeptracker," 2016. [Online]. Available: <http://www.sleepimage.com/>
- [23] M. Bsoul, H. Minn, and L. Tamil, "Apnea MedAssist: Real-time sleep apnea monitor using single-lead ECG," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 3, pp. 416–427, May 2011.
- [24] J. M. Kelly, R. E. Strecker, and M. T. Bianchi, "Recent developments in home sleep-monitoring devices," *ISRN Neurology*, vol. 2012, 2012, Art no. 768794.
- [25] M. Jarvis and P. Mitra, "Apnea patients characterized by 0.02 Hz peak in the multitaper spectrogram of electrocardiogram signals," in *Proc. IEEE Comput. Cardiology*, 2000, pp. 769–772.
- [26] B. Raymond, R. Cayton, R. Bates, and M. Chappell, "Screening for obstructive sleep apnoea based on the electrocardiogram—the CinC challenge," in *Proc. IEEE Comput. Cardiology*, 2000, pp. 267–270.
- [27] P. De Chazal, C. Heneghan, E. Sheridan, R. Reilly, P. Nolan, and M. Malley, "Automatic classification of sleep apnea epochs using the electrocardiogram," in *Proc. IEEE Comput. Cardiology*, 2000, pp. 745–748.
- [28] J. McNames and A. Fraser, "Obstructive sleep apnea classification based on spectrogram patterns in the electrocardiogram," in *Proc. IEEE Comput. Cardiology*, 2000, pp. 749–752.
- [29] P. Stein and P. Domitovich, "Detecting OSAHS from patterns seen on heart-rate tachograms," in *Proc. IEEE Comput. Cardiology*, 2000, pp. 271–274.
- [30] J. Mietus, C. Peng, P. C. Ivanov, and A. L. Goldberger, "Detection of obstructive sleep apnea from cardiac interbeat interval time series," in *Proc. IEEE Comput. Cardiology*, 2000, pp. 753–756.
- [31] Z. Shinar, A. Baharav, and S. Akselrod, "Obstructive sleep apnea detection based on electrocardiogram analysis," in *Proc. IEEE Comput. Cardiology*, 2000, pp. 757–760.
- [32] M. Drinnan, J. Allen, P. Langley, and A. Murray, "Detection of sleep apnoea from frequency analysis of heart rate variability," in *Proc. IEEE Comput. Cardiology*, 2000, pp. 259–262.
- [33] C. Maier, M. Bauch, and H. Dickhaus, "Recognition and quantification of sleep apnea by analysis of heart rate variability parameters," in *Proc. IEEE Comput. Cardiology*, 2000, pp. 741–744.

- [34] M. Schrader, C. Zywiets, V. Von Einem, B. Widiger, and G. Joseph, "Detection of sleep apnea in single channel ECGs from the PhysioNet data base," in *Proc. IEEE Comput. Cardiology*, 2000, pp. 263–266.
- [35] F. Ng, I. Garcia, P. Gomis, A. La Cruz, G. Passariello, and F. Mora, "Bayesian hierarchical model with wavelet transform coefficients of the ECG in obstructive sleep apnea screening," in *Proc. IEEE Comput. Cardiology*, 2000, pp. 275–278.
- [36] P. De Chazal, C. Heneghan, E. Sheridan, R. Reilly, P. Nolan, and M. O'Malley, "Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 6, pp. 686–696, Jun. 2003.
- [37] A. M. da Silva Pinho, N. Pombo, and N. M. Garcia, "Sleep apnea detection using a feed-forward neural network on ECG signal," in *Proc. IEEE 18th Int. Conf. E-Health Netw., Appl. Serv.*, 2016, pp. 1–6.
- [38] S.-H. Fan, C.-C. Chou, W.-C. Chen, and W.-C. Fang, "Real-time obstructive sleep apnea detection from frequency analysis of EDR and HRV using lomb periodogram," in *Proc. IEEE 37th Annu Int. Conf. Eng. Med. Biol. Soc.*, 2015, pp. 5989–5992.
- [39] J. D. Prathap, E. Rangan, and R. K. Pathinarupothi, "Real-time and offline techniques for identifying obstructive sleep apnea patients," in *Proc. IEEE Int. Conf. Comput. Intell. Comput. Res.*, 2016, pp. 1–4.
- [40] P. De Chazal and C. Heneghan, "Apparatus for detecting sleep apnea using electrocardiogram signals," U.S. Patent 7,025,729, 2006.
- [41] R. K. Pathinarupothi, R. Vinaykumar, E. Rangan, E. Gopalakrishnan, and K. Soman, "Instantaneous heart rate as a robust feature for sleep apnea severity detection using deep learning," in *Proc. IEEE Int. Conf. Biomed. Health Inform.*, 2017, pp. 293–296.
- [42] G. Surrel, F. Rincón, S. Murali, and D. Atienza, "Low-power wearable system for real-time screening of obstructive sleep apnea," in *Proc. IEEE VLSI (ISVLSI) Comput. Soc. Annu. Symp.*, 2016, pp. 230–235.
- [43] D. Sopic, A. Aminifar, A. Aminifar, and D. Atienza, "Real-time classification technique for early detection and prevention of myocardial infarction on wearable devices," in *Proc. 13th IEEE Biomed. Circuits Syst. Conf.*, 2017, pp. 1–4.
- [44] D. Sopic, E. D. Giovanni, A. Aminifar, and D. Atienza, "A hierarchical cardiac rhythm classification methodology based on electrocardiogram fiducial points," in *Proc. IEEE Comput. Cardiology*, 2017, pp. 1–4.
- [45] E. D. Giovanni, A. Aminifar, A. Luca, S. Yazdani, J.-M. Vesin, and D. Atienza, "A patient-specific methodology for prediction of paroxysmal atrial fibrillation onset," in *Proc. Comput. Cardiology*, 2017, pp. 1–4.
- [46] "ECG electrode for sensitive skin | ambu BlueSensor VLC," 2014. [Online]. Available: http://www.ambu.com/corp/products/patient_monitoring_and_diagnostics/product/ambu%20AE_bluSENSOR_vlc-prd837.aspx
- [47] "STM321151rd - ultra-low-power ARM cortex-m3 MCU with 384 kbytes flash, 32 MHz CPU, USB, 3xop-amp - STMicroelectronics," 2017. [Online]. Available: <http://www.st.com/en/microcontrollers/stm321151rd.html>
- [48] J. G. Webster, "Reducing motion artifacts and interference in biopotential recording," *IEEE Trans. Biomed. Eng.*, vol. 31-BME, no. 12, pp. 823–826, Dec. 1984.
- [49] J. C. Huhta and J. G. Webster, "60-hz interference in electrocardiography," *IEEE Trans. Biomed. Eng.*, vol. 20-BME, no. 2, pp. 91–101, Mar. 1973.
- [50] Y. Sun, K. L. Chan, and S. M. Krishnan, "ECG signal conditioning by morphological filtering," *Comput. Biol. Med.*, vol. 32, no. 6, pp. 465–479, 2002.
- [51] R. Braojos, G. Ansaloni, D. Atienza, and F. J. Rincón, "Embedded real-time ECG delineation methods: A comparative evaluation," in *Proc. IEEE 12th Int. Conf. Bioinform. Bioeng.*, 2012, pp. 99–104.
- [52] N. Boichat, N. Khaled, F. Rincon, and D. Atienza, "Wavelet-based ECG delineation on a wearable embedded sensor platform," in *Proc. IEEE 6th Int. Workshop Wearable Implantable Body Sensor Netw.*, 2009, pp. 256–261.
- [53] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vanderghyest, "Compressed sensing for real-time energy-efficient ECG compression on wireless body sensor nodes," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 9, pp. 2456–2466, Sep. 2011.
- [54] F. J. Rincón *et al.*, "Implementation of an automated ECG-based diagnosis for a wireless body sensor platform," in *Proc. Int. Conf. Biomed. Electron Devices*, 2009, pp. 88–96.
- [55] "Cc2640r2f (active)," 2016. [Online]. Available: <http://www.ti.com/product/cc2640r2f/datasheet>
- [56] M. Crepaldi, C. Li, K. Dronson, J. Fernandes, and P. Kinget, "An ultra-low-power interference-robust IR-UWB transceiver chipset using self-synchronizing OOK modulation," in *Proc. IEEE Int. Solid-State Circuits Conf. Digest Tech. Papers*, 2010, pp. 226–227.
- [57] S. S. Basu, L. G. Duch, R. Braojos Lopez, G. Ansaloni, L. Pozzi, and D. A. Alonso, "An inexact ultra-low power bio-signal processing architecture with lightweight error recovery," in *Proc. CODES+ ISSS: Int. Conf. Hardware/Softw. Codesign Syst. Synthesis*, 2017, Art no. 159.
- [58] J. Tang, S. Alelyani, and H. Liu, "Feature selection for classification: A review," *Data Classification, Algorithms Appl.*, CRC Press, p. 37, 2014. [Online]. Available: http://eprints.kku.edu.sa/17011/feature_selection_for_classification.pdf
- [59] E. A. Ashley and J. Niebauer, *Cardiology Explained*. London, U.K.: Remedica, 2004.
- [60] T. Penzel, J. McNames, P. De Chazal, B. Raymond, A. Murray, and G. Moody, "Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings," *Med. Biological Eng. Comput.*, vol. 40, no. 4, pp. 402–407, 2002.
- [61] B. Xie and H. Minn, "Real-time sleep apnea detection by classifier combination," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 469–477, May 2012.
- [62] T. K. Ho, "Random decision forests," in *Proc. 3rd IEEE Int. Conf. Document Anal. Recognit.*, 1995, vol. 1, pp. 278–282.
- [63] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
- [64] G. Clifford and L. Tarassenko, "Quantifying errors in spectral estimates of HRV due to beat replacement and resampling," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 4, pp. 630–638, Apr. 2005.
- [65] M. Rienzner, "Find Outliers with Thompson Tau," 2010. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/27553-find-outliers-with-thompson-tau/content/find_outliers_Thompson.m
- [66] C. A. R. Hoare, "Algorithm 65: find," *Commun. ACM*, vol. 4, no. 7, pp. 321–322, 1961.
- [67] Student, "The probable error of a mean," *Biometrika*, vol. 6, no. 1, pp. 1–25, 1908.
- [68] B. Welford, "Note on a method for calculating corrected sums of squares and products," *Technometrics*, vol. 4, no. 3, pp. 419–420, 1962.
- [69] G. Karakonstantis, A. Sankaranarayanan, M. M. Sabry, D. Atienza, and A. Burg, "A quality-scalable and energy-efficient approach for spectral analysis of heart rate variability," in *Proc. IEEE Design, Autom. Test Eur. Conf. Exhib.*, 2014, pp. 1–6.
- [70] M.-A. Parseval, "Mémoire sur les séries et sur l'intégration complète d'une équation aux différences partielles linéaires du second ordre, à coefficients constants," *Mém. prés. par divers savants, Acad. des Sci., Paris.(1)*, vol. 1, pp. 638–648, 1806.
- [71] S. Butterworth, "On the theory of filter amplifiers," *Wireless Eng.*, vol. 7, no. 6, pp. 536–541, 1930.
- [72] T. Penzel *et al.*, "Apnea-ECG Database," 2000. [Online]. Available: <http://dx.doi.org/10.13026/C23W2R>
- [73] G. Moody, R. Mark, A. Goldberger, and T. Penzel, "Stimulating rapid research advances via focused competition: The computers in cardiology challenge 2000," in *Proc. IEEE Comput. Cardiology*, 2000, pp. 207–210.
- [74] T. Instruments. Complete low power integrated analog front end for eeg applications, 2012. [Online]. Available: <http://www.ti.com/product/ADS1191>
- [75] T. InvenSense. Mpu-6000 motion sensor, 2012. [Online]. Available: <https://store.invensense.com/ProductDetail/MPU6000-InvenSense-Inc/420595/>
- [76] S. Microelectronics. Ultra-low-power arm cortex-m3 MCU with 384 kbytes flash, 32 MHz CPU, USB, 3xop-amp. [Online]. Available: <http://www.st.com/en/microcontrollers/stm321151rd.html>
- [77] "NRF8001 bluetooth low energy connectivity IC," 2013. [Online]. Available: <https://www.nordicsemi.com/eng/Products/Bluetooth-low-energy/nRF8001>
- [78] C. Orphanidou, T. Bonnici, P. Charlton, D. Clifton, D. Vallance, and L. Tarassenko, "Signal-quality indices for the electrocardiogram and photoplethysmogram: Derivation and applications to wireless monitoring," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 3, pp. 832–838, May 2015.



Grégoire Surrel received the M.Sc. degree in electrical engineering from CPE Lyon, Lyon, France, in 2014. He is currently working toward the Ph.D. degree at the Embedded Systems Laboratory of Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland with Prof. Atienza. During a scientific collaboration with the Integrated System Laboratory in EPFL, he developed a mobile biosignal collection system for noninvasive data collection. His current research includes biosignal acquisition and optimized processing on low-power wearable systems for online health-care screening and diagnostics.



Amir Aminifar received the Ph.D. degree from the Swedish National Computer Science Graduate School, Linköping University, Linköping, Sweden, in 2016. During 2014–2015, he visited the Cyber-Physical Systems Laboratory of the University of California, Los Angeles, CA, USA, and the Real-Time Systems Laboratory, Scuola Superiore, Sant'Anna, Italy. He is currently a Research Scientist with the Embedded Systems Laboratory of Swiss Federal Institute of Technology, Lausanne, Switzerland. His current research interests include

centered around mobile health and smart wearable technologies for early detection/prediction of pathological health conditions.



Francisco Rincón received the M.Sc. and Ph.D. degrees in computer science and engineering from the Complutense University of Madrid, Madrid, Spain, in 2006 and 2012, respectively. He is currently the CTO and cofounder of SmartCardia, Lausanne, Switzerland. He is coauthor of 4 U.S. patents and more than 30 publications in leading conferences and journals. His current research interests include embedded signal processing, on-device machine learning, and energy-aware optimizations for wearable medical devices.



Srinivasan Murali received the M.Sc. and Ph.D. degrees in electrical engineering from Stanford University, New Delhi, India, in 2007. He is currently the CEO and cofounder of SmartCardia, Lausanne, Switzerland, a wearable medical device company. Previously, he was a Program Manager of sensor systems with IMEC and was a cofounder and CTO of iNoCs. He has authored/coauthored a book, several patents, book chapters, and around 60 publications in leading conferences and journals (around 5000 citations, with h-index of 29 and i10-index of 42). His

research interests include wearable devices and IoT, on-device machine learning and biosignal analysis, interconnect design for systems on chips, with particular emphasis on developing CAD tools and design methods for networks on chips, and also include thermal modeling and reliability of multicore systems. He is the recipient of the EDAA Outstanding Dissertation Award in 2007 for his work on the interconnect architecture design. He was the recipient of best Paper Award at the DATE in 2005 conference, and a best paper nomination of International Conference on Control, Automation and Diagnosis in 2006. One of his papers has also been selected as one of The Most Influential Papers of 10 Years DATE.



David Atienza (M'05–SM'13–F'16) received the M.Sc. and Ph.D. degrees in computer science and engineering from the Complutense University of Madrid, Madrid, Spain, and Inter-University Microelectronics Center, Leuven, Belgium, in 2001 and 2005, respectively. He is currently an Associate Professor in electrical and computer engineering, and the Director of the Embedded Systems Laboratory with the Swiss Federal Institute of Technology Lausanne, Lausanne, Switzerland. His research interests include system-level design methodologies for high-

performance multiprocessor system-on-chip (MPSoC) and low-power embedded systems, including new 2-D/3-D thermal-aware design for MPSoCs and many-core servers, ultralow-power system architectures for wireless body sensor nodes, HW/SW reconfigurable systems, dynamic memory optimizations, and network-on-chip design. He is a coauthor of more than 250 publications in peer-reviewed international journals and conferences, several book chapters, and seven U.S. patents in these fields. He was the recipient of several best paper awards, and is (or has been) an Associate Editor of IEEE TRANSACTIONS ON COMPUTERS, the IEEE DESIGN AND TEST OF COMPUTERS, the IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS, the IEEE TRANSACTIONS ON SUSTAINABLE COMPUTING, and Elsevier *Integration*. He was the Technical Programme Chair of IEEE/ACM DATE 2015 and General Programme Chair of IEEE/ACM DATE 2017. He was the recipient of several best paper awards, and also an ERC Consolidator Grant in 2016, the IEEE CEDA Early Career Award in 2013, the ACM SIGDA Outstanding New Faculty Award in 2012, a Faculty Award from Sun Laboratory for Oracle in 2011, and he was a Distinguished Lecturer (period 2014–2015) of IEEE Circuits and Systems Society. He is an ACM Distinguished Member.