IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 15, 2019, accepted August 4, 2019, date of publication September 5, 2019, date of current version September 20, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2939749

Entropy Analysis of Acoustic Signals Recorded With a Smartphone for Detecting Apneas and Hypopneas: A Comparison With a Commercial System for Home Sleep Apnea Diagnosis

YOLANDA CASTILLO-ESCARIO[1,2,3], IGNASI FERRER-LLUIS[1,2,3], JOSEP MARIA MONTSERRAT[4,5], AND RAIMON JANÉ[1,2,3], (Senior Member, IEEE)

[1]Institute for Bioengineering of Catalonia, Barcelona Institute of Science and Technology, 08028 Barcelona, Spain
[2]Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, 28029 Madrid, Spain
[3]Department of Automatic Control, Universitat Politècnica de Catalunya-Barcelona Tech, 08028 Barcelona, Spain
[4]Sleep Lab, Pneumology Service, Hospital Clínic de Barcelona, 08036 Barcelona, Spain
[5]Centro de Investigación Biomédica en Red de Enfermedades Respiratorias, 28029 Madrid, Spain

Corresponding author: Yolanda Castillo-Escario (ycastillo@ibecbarcelona.eu)

This work was supported in part by ‘‘"La Caixa’’ Foundation (ID 100010434); the fellowship codes are LCF/BQ/ES18/11670019 and LCF/BQ/IN17/11620029, in part by the European Union’s Horizon 2020 Research and Innovation Program under the Marie Sklodowska-Curie under Grant 713673, in part by the CERCA Program/Generalitat de Catalunya, in part by the Secretaria d’Universitats i Recerca de la Generalitat de Catalunya under Grant GRC 2017 SGR 01770, in part by the Spanish Ministry of Economy and Competitiveness under Grant DPI2015-68820-R MINECO/FEDER and Grant RTI2018-098472-B-I00 MCIU/AEI/FEDER, UE, and in part by the Instituto de Salud Carlos III under Grant FIS PI17/01068.

ABSTRACT Obstructive sleep apnea (OSA) is a prevalent disease, but most patients remain undiagnosed and untreated. Here we propose analyzing smartphone audio signals for screening OSA patients at home. Our objectives were to: (1) develop an algorithm for detecting silence events and classifying them into apneas or hypopneas; (2) evaluate the performance of this system; and (3) compare the information provided with a type 3 portable sleep monitor, based mainly on nasal airflow. Overnight signals were acquired simultaneously by both systems in 13 subjects (3 healthy subjects and 10 OSA patients). The sample entropy of audio signals was used to identify apnea/hypopnea events. The apnea-hypopnea indices predicted by the two systems presented a very high degree of concordance and the smartphone correctly detected and stratified all the OSA patients. An event-by-event comparison demonstrated good agreement between silence events and apnea/hypopnea events in the reference system (Sensitivity = 76%, Positive Predictive Value = 82%). Most apneas were detected (89%), but not so many hypopneas (61%). We observed that many hypopneas were accompanied by snoring, so there was no sound reduction. The apnea/hypopnea classification accuracy was 70%, but most discrepancies resulted from the inability of the nasal cannula of the reference device to record oral breathing. We provided a spectral characterization of oral and nasal breathing to correct this effect, and the classification accuracy increased to 82%. This novel knowledge from acoustic signals may be of great interest for clinical practice to develop new non-invasive techniques for screening and monitoring OSA patients at home.

INDEX TERMS Acoustics, biomedical signal processing, mHealth, monitoring, sleep apnea, smartphone.

I. INTRODUCTION

Sleeping, like breathing, is an action that we undertake throughout our lives. We spend more than 25% of our time asleep, and this is strictly necessary, since sleep is an organism’s natural state of rest and self-regulation. However, several diseases can affect sleep quality, and this can translate into symptoms of varying severity.

Obstructive sleep apnea (OSA) is one of the most common and serious sleep disorders. OSA is characterized by repetitive episodes of total or partial airflow reduction during sleep, produced by upper airway obstruction. These breathing

The associate editor coordinating the review of this manuscript and approving it for publication was Vishal Srivastava.

disturbances lead to hypoxia and microarousals, seriously affecting the patient's quality of sleep. Sleep fragmentation produces fatigue and daytime sleepiness, with the subsequent risk of domestic, work-related, and traffic accidents. In addition, OSA has long been associated with increased cardiovascular and cerebrovascular morbidity and mortality [1]–[3].

OSA is a very prevalent disease, especially in the elderly and obese population. A recent systematic review suggests that the overall population prevalence ranges from 9% to 38%, being higher in men, and increases up to 90% in men and 78% in women in some elderly groups [4]. However, despite these overwhelming statistics and the serious consequences of OSA, most patients remain undiagnosed and untreated [5]. This constitutes an important social, health, and economic burden.

The gold-standard for OSA diagnosis is nocturnal polysomnography (PSG), which is a type 1 monitoring system. In PSG studies, many physiological signals are recorded from the patient, in an overnight stay at hospital. The information from these signals is combined to extract a diagnostic index. The most common is the Apnea/Hypopnea Index (AHI), which is calculated as the number of apneas and hypopneas per hour of sleep. According to the American Academy of Sleep Medicine (AASM), an apnea is defined as a $\geq$90% flow reduction for $\geq$10 seconds [6]. The AASM guidelines provide two definitions for hypopnea: the ''recommended'' one requires a $\geq$30% flow reduction for $\geq$10 seconds associated with $\geq$4% oxygen desaturation; the ''alternative'' version involves a $\geq$50% flow reduction followed by a $\geq$3% oxygen desaturation or an arousal [6]. Patients are categorized according to their AHI values as follows: normal (AHI $<$ 5); mild OSA (5 $\leq$ AHI $<$ 15); moderate OSA (15 $\leq$ AHI $<$ 30); and severe OSA (AHI $\geq$ 30) [6].

Nevertheless, PSG has important limitations, ranging from its elevated complexity and cost, to the high level of discomfort for patients, who have to spend a night in hospital and may find it difficult to sleep attached to so many sensors and wires. Medical doctors still evaluate the results by visual inspection, which is tiring and time-consuming. Moreover, the diagnosis is based on a single night, without considering the possible variability between nights. On the other hand, as resources are limited and it takes an entire night to diagnose each patient, PSG cannot be applied to all patients who might need it. For these reasons, there is an increasing need for alternative diagnostic devices.

Novel systems that are being proposed for detecting OSA are usually simple, portable devices that can be used at home. Instead of recording many signals like PSG, they tend to use inexpensive unobtrusive sensors to record just a single or a reduced number of signals that have a high diagnostic power. Some of the channels that have been tested for these approaches include nasal airflow, thoracic effort, oxygen saturation, actigraphy, and acoustic signals [7].

The most direct approach for detecting OSA involves measuring nasal airflow with nasal pressure transducers, such as nasal cannulas. This appears to be the easiest and most reliable way of detecting apnea and hypopnea events, since they are defined as airflow reductions. However, patients can also breathe through their mouths, and nasal cannulas cannot record this kind of flow; for this reason they overestimate breathing disturbance when oral breathing takes place [8]. Healthy humans preferentially breathe via their noses, but nasal obstruction leads to increased oral breathing, and this condition is more prevalent in subjects with OSA [9]. It has also been reported that patients with OSA spend more time breathing orally, even in the absence of nasal obstruction [10]. Other studies found that oral breathing induces OSA or worsens it [11], even reducing the effectiveness of CPAP or oral appliance therapies [12], [13]. In this context, the inability to capture oral breathing limits the diagnostic accuracy of nasal airflow sensors.

Pulse oximetry is a simple non-invasive method for measuring blood oxygen saturation (SpO_2). Local reductions in SpO_2 (i.e., desaturation) followed by reoxygenation are typically associated with apnea and hypopnea episodes [14]. For this reason, pulse oximetry has also been used as a primary signal for discriminating OSA [15]–[17].

Another emerging approach for diagnosing OSA involves analyzing acoustic snoring signals. These signals can be acquired using microphones, which are simple and low-cost sensors, available in a wide range of models and solutions. Apneas and hypopneas can be identified in audio recordings as an absence or reduction of sound. Moreover, the characteristics of breathing and snoring sounds provide valuable information about the mechanisms and degree of upper airway obstruction [18]. Snoring is one of the most common and earliest symptoms of OSA, and snoring sounds differ in healthy and sleep apnea patients, as reported by Fiz *et al.* [19]. For this reason, the acoustic analysis of snoring has been used to extract information for detecting and classifying OSA patients [20]. Some prototypes have also been proposed to record respiratory sounds during sleep, such as the single-channel device developed by Jané *et al.* [21], which automatically analyzed snoring to screen patients with OSA.

In recent years, smartphones have become ever more present in our daily lives. This enables us to take advantage of their multiple capabilities and turn them into suitable solutions for mobile health (mHealth) monitoring. Many attempts have been made to monitor health and disturbed sleep using smartphone apps [22]. However, most focus on sleep quality and activity monitoring rather than OSA detection. Garde *et al.* [23] used SpO_2 and pulse rate variability, obtained using a pulse oximeter integrated with a smartphone, to detect OSA events in children. Nakano *et al.* [24] used the built-in microphone of a smartphone to record ambient sound and quantify snoring and OSA severity. Recently, our group worked with smartphone microphones and accelerometers to characterize breathing, snoring and OSA-related patterns [25], [26]. Other studies have combined multiple channels from the smartphone or external sensors to extract

features and obtain OSA risk indices [27]–[29]. Nevertheless, despite all these efforts, more validation studies are needed to accurately assess the clinical feasibility of these screening applications [22].

In this work, we propose an mHealth system based on the built-in microphone of a smartphone, to record overnight breathing acoustic signals and use them for OSA screening. We present a method for analyzing these signals and detecting apnea and hypopnea events. We compare the results with a commercially available type 3 portable device for diagnosing OSA. We have three main objectives: (1) to develop an automatic algorithm for detecting silence events and classifying them into apneas or hypopneas; (2) to evaluate the performance of the smartphone system for detecting and monitoring OSA patients; and (3) to compare the information provided by audio signals from the smartphone and nasal airflow signals from a portable device, in terms of apnea, hypopnea, and breathing events.

II. MATERIALS AND METHODS

A. DATABASE AND ACQUISITION PROTOCOL

To compare the performance of the proposed smartphone-based approach with a portable commercial device, full-night recordings were acquired simultaneously using the two systems in the patients' homes. All the experiments were approved by the Hospital Clínic Ethics Committee.

The commercial equipment for sleep apnea diagnosis that was taken as a reference for this study was ApneaLink™ Air (ResMed Inc, San Diego, California, USA). The smartphone that we selected was a Samsung Galaxy S5, since it is a mid-range product but has a high-quality microphone, as reported in previous studies by our group [25], [26].

Thirteen participants were enrolled in the study. Table 1 provides their anthropometric description, and OSA severity according to ApneaLink diagnosis. As described in the table, the sample comprised 8 men and 5 women, aged 25 to 83; it included 3 healthy subjects and 10 subjects diagnosed with sleep apnea (3 mild OSA, 4 moderate OSA, and 3 severe OSA). We will refer to this database as DB1.

During the full-night recordings at home, the smartphone was attached to each subject's chest with an elastic band, in the position suggested by Nakano *et al.* [24]. The ApneaLink device was fixed just below, over the sternum (Fig. 1). ApneaLink measured respiratory flow through a nasal cannula, thoracic movement using a belt connected to an effort sensor, and SpO_2 through a wired fingertip pulse oximeter. The sampling frequency (fs) was 100 Hz for airflow, 10 Hz for thoracic effort and 1 Hz for SpO_2. The built-in microphone on the smartphone recorded overnight acoustic signals, with a sampling rate of 48 kHz. Its embedded triaxial accelerometer (fs = 200 Hz) was also used to track the subjects' movements and position changes.

To simplify the acquisition process, data recording was managed using the Android app "Automate", which automatically launched the acquisition apps when the phone booted up. The apps used were Easy Voice Recorder for audio signals, and Sensors Logger for accelerometry. The data were automatically saved in the smartphone's local memory, in .wav and.txt formats, respectively. The internal memory was enough to store a full night's acquisition.

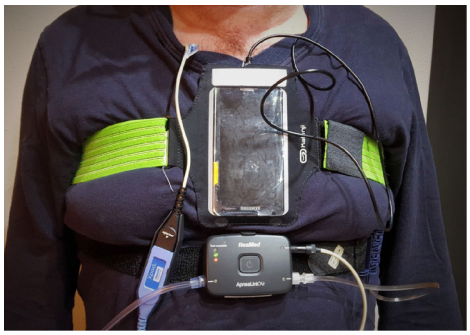

FIGURE 1. Position of smartphone and ApneaLink device on the subject's chest.

TABLE 1. Database anthropometric description.

Subject	Gender	Age	BMI[a]	OSA Severity
1	Male	62	27.8	Moderate
2	Female	58	29	Moderate
3	Male	25	25.9	Mild
4	Male	61	31.4	Severe
5	Female	24	21.8	Healthy
6	Female	25	20	Healthy
7	Male	55	27.7	Moderate
8	Male	56	29.8	Moderate
9	Male	60	22.9	Mild
10	Female	25	33.8	Healthy
11	Male	58	31.3	Severe
12	Female	83	26.5	Severe
13	Male	38	26.6	Mild

Colors indicate OSA severity: healthy (green), mild (yellow), moderate (orange), and severe (red).
[a]BMI: Body mass index (kg/m^2)

B. SIGNAL PROCESSING AND ANALYSIS

The ApneaLink signals were automatically analyzed by its proprietary software to identify apnea, hypopnea, desaturation (≥3%), and snoring events. However, the results of this analysis were also manually reviewed by an expert in sleep medicine from Hospital Clínic de Barcelona, following the same criteria used in PSG. These reviewed event labels were used as the ground truth for this study. The ApneaLink signals were also exported in European Data Format (.EDF) for further analysis.

Smartphone signals were processed and analyzed offline using custom-made functions and algorithms in the MATLAB®programming environment (r2018a, Mathworks Inc.). Firstly, they were manually synchronized to the reference system. Regions that had not been analyzed in ApneaLink, due to artifacts or low signal magnitude, were discarded.

Movement artifacts and position changes were detected from the smartphone accelerometry [30] and excluded from

the analysis, since they also produced sound artifacts. Finally, the position signal, also extracted from accelerometry [30], allowed us to discard the segments of the night where the subject was standing. The final valid length of the signals was around 4-5 hours for all subjects.

The audio signals were downsampled to 5 kHz, applying an anti-aliasing low-pass filter with a cut-off frequency of 2.5 kHz. The 50 Hz power-line noise and harmonics were removed using a Notch filter. The signals were also forward-backward filtered with an 8^{th} order Butterworth band-pass filter with cut-off frequencies of 70 Hz and 2 kHz, to reduce cardiac and high-frequency noise, while retaining the main band of interest for breathing and snoring sounds [21]. Lastly, the signals were normalized to their maximum amplitude.

C. DETECTING SILENCE EVENTS

1) SAMPLE ENTROPY

The automatic detector that we propose is based on the calculation of the sample entropy (SampEn), as a measure of signal complexity. Our algorithm is aimed at detecting silence events (SEv), which can correspond to either apnea or hypopnea events (AHEv). This detector is an improved version of a preliminary algorithm presented in [31].

SampEn is a measure of time-series regularity, or complexity, and is a robust estimator suitable for short and noisy physiological signals [32]. Let N be the number of data points in the time series, m the embedding dimension (i.e., the length of the subsets to be compared), and r a tolerance parameter. Then, SampEn(m,r,N) is defined as the negative natural logarithm of the conditional probability that, in a dataset of length N, two sequences that are similar for m samples within a tolerance r remain similar for $m + 1$ samples [32], [33].

Lower SampEn values indicate a greater self-similarity and predictability in the time-series. Snoring and breathing present more complex patterns than silence regions, hence producing higher SampEn values. In addition, if r is fixed for a certain segment, then SampEn also depends on the signal amplitude and can be used as an envelope estimator [31].

2) ALGORITHM FOR DETECTING SILENCE EVENTS

We analyzed the regions of the audio signals that were followed by desaturations (according to ApneaLink labels). In this way, we used SpO_2 to guide our detector towards the regions of interest and thus reduce the computational cost and false alarm rate. In this case, SpO_2 was obtained using a non-smartphone-based device, because we wanted to analyze the same windows as this reference system. Specifically, we built signal windows starting 60 s prior to the start of each desaturation and ending when the desaturation finished, as shown in Fig. 2. Overlapping segments were concatenated, up to a maximum length of 10 minutes.

The step-by-step algorithm that was applied to each segment is illustrated in Fig. 2. The method is the following:

1. Calculate the SampEn of the audio signal in that segment, using 0.75 s windows with 50% overlap, and SampEn parameters $m = 2$ and r equal to the standard deviation of the segment.
2. Subtract from SampEn the value of its 25^{th} percentile to remove any possible offset. During breathing, about half the time is spent in pauses between respiratory events. Therefore, the 25^{th} percentile should correspond to a value in a silence region.
3. Compute the threshold for SEv as the maximum value between the 60^{th} percentile of SampEn and a fixed value of 0.01. In our previous work [31], we used a fixed threshold, but the 60^{th} percentile allows better adaptation to the signal-to-noise ratio (SNR) and the characteristics of each window. However, in regions with many apneas, in severe patients, this may be insufficient. For this reason, we used the maximum of these two values.
4. Find all SampEn points below this threshold.
5. Select regions longer than 6 s and shorter than 100 s as potential SEv. AHEv are defined as reductions in airflow longer than 10 s [6], but these are usually shorter in audio signals, due to the high sensitivity to low-intensity sounds. Thus, we were less restrictive with the minimum duration for SEv, as in previous work on acoustic signals [21]. The upper limit of 100 s is to avoid false long events in regions where the signal magnitude is too low.
6. Merge events that are less than 3 seconds apart.
7. Directly associate the detected events to desaturations and discard unpaired events. To do this, we considered that an event was followed by a desaturation if the latter met the following two conditions:

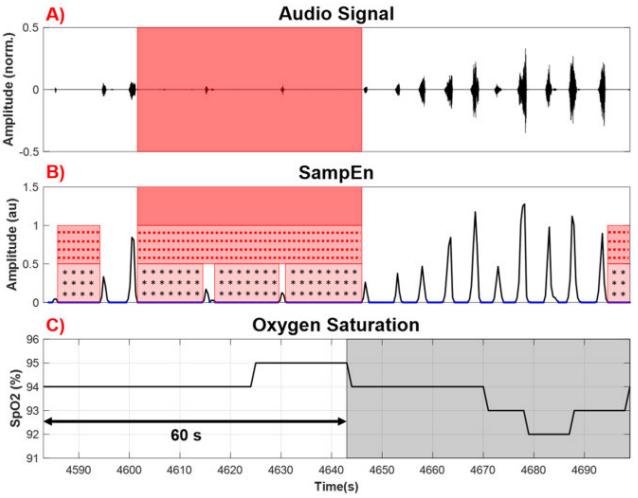

FIGURE 2. Depiction of the algorithm for detecting silence events. From audio signals (A), SampEn (B) is calculated in windows starting 60s before the desaturation (C). The points below a certain threshold are identified (shown in blue in B) and regions > 6s are retained (boxes with black asterisks). Close events are merged (boxes with red dotted lines). Lastly, events which are not directly associated with the desaturation are removed (first and last events in this example). The remaining episodes are the final silence events identified by the automatic detector (plain red box shown in A and B panels).

a. It started in the time interval [t_ini_{EV} − 10 s, t_end_{EV} + 20 s]
b. It reached a local minimum, representing a drop ≥3% in SpO_2, within the time interval [t_ini_{EV}+5 s, t_end_{EV} + 45 s]

where t_ini_{EV} and t_end_{EV} are the initial and final times of the silence event, respectively.
The values for the previous parameters were empirically optimized after analyzing a previous polysomnographic database containing more than 2000 events.

3) COMPARISON WITH THE REFERENCE SYSTEM

To evaluate the ability of the proposed system for screening sleep apnea, we first counted the number of SEv that had been detected and calculated the predicted AHI for each patient. We plotted these values against the AHIs reported by ApneaLink, and computed the concordance correlation coefficient, ρ_c [34]. For two paired variables, x_n and y_n, ρ_c can be calculated as:

$$\rho_c = \frac{2\sigma_{xy}}{\sigma_x^2 + \sigma_y^2 + (\overline{x} - \overline{y})^2} \tag{1}$$

where σ_{xy} is the covariance of x_n and y_n, σ_x^2 and σ_y^2 are the variances of x_n and y_n, and $\overline{x}$ and $\overline{y}$ are the means of x_n and y_n, respectively. In our case, x_n and y_n were vectors of length 13 (the number of participants in the study) containing the AHIs predicted for each subject by ApneaLink and the smartphone system, respectively.

After that, we evaluated the performance of our detector by comparing it to ApneaLink in terms of event-by-event agreement. We took all the events from the subjects with AHI ≥ $5h^{-1}$ and considered that a SEv matched an event in the reference system if these two events overlapped. Under this premise, it was possible for more than one SEv to match a single event from ApneaLink, or more than one ApneaLink event to match a single SEv. To deal with this issue, we computed the number of true positives (TP), false positives (FP), and false negatives (FN) with respect to the reference. This means that ApneaLink events were counted as TP if they overlapped any SEv, and as FN if they did not; while any SEv that did not match an ApneaLink event was counted as FP. Finally, these values were used to calculate sensitivity (Se) and positive predictive value (PPV):

$$Se(\%) = 100\frac{TP}{(TP + FN)} \tag{2}$$

$$PPV(\%) = 100\frac{TP}{(TP + FP)} \tag{3}$$

Specificity (Sp) and negative predictive value (NPV) cannot be calculated in this event-by-event comparison, since they are defined as:

$$Sp(\%) = 100\frac{TN}{(TN + FP)} \tag{4}$$

$$NPV(\%) = 100\frac{TN}{(TN + FN)} \tag{5}$$

but the concept of 'true negative' (TN) does not exist in this case, because it would signify an event that is not found in either system.

We also calculated the percentages of each type of event according to ApneaLink (apnea/hypopnea) that we detected as silence events, i.e.,:

$$\%Apn = 100 \cdot \frac{Number\ of\ apneas\ matching\ a\ SEv}{Total\ number\ of\ apneas} \tag{6}$$

$$\%Hpn = 100 \cdot \frac{Number\ of\ hypopneas\ matching\ a\ SEv}{Total\ number\ of\ hypopneas} \tag{7}$$

D. CLASSIFYING APNEAS/HYPOPNEAS

After comparing the automatically detected SEv with the events in the reference system, we also tried to classify them into apneas or hypopneas. For this, we took the TP events (those found in both the audio signals and ApneaLink) and developed an algorithm for zooming in on SEv and searching for low-intensity respiratory sounds (LRS). If this residual breathing activity was present, then the event should be classified as a hypopnea, while if there was no breathing at all, it would be labeled as an apnea. However, working at this level of magnification, the resolution and SNR is limited, and the task of discriminating between LRS and artifacts becomes more difficult. For this reason, several filtering and refinement steps are required to optimize the conditions and improve the performance of the classification. We implemented a combination of techniques in the time and time-frequency domains, as described below and shown in Fig. 3. We worked separately on the acoustic signal segments corresponding to each SEv.

1) ALGORITHM FOR CLASSIFYING APNEAS/HYPOPNEAS

Each segment was first filtered with the Wiener noise suppressor proposed in [35], which uses the two-step noise reduction technique (TSNR) followed by a method called harmonic regeneration noise reduction (HRNR). This filter refines the *a priori* SNR estimated by a decision-directed method for computing the optimal spectral gain for filtering, removing reverberation effects and avoiding harmonic distortion [35]. A noise reference is required as an input for this filter. For each SEv, we used the 1-second window of the signal with the lowest root-mean-square (RMS) value as the noise reference.

An envelope based on local maxima and minima was extracted from these filtered signals. Specifically, we split the signal into 10 ms windows and computed the difference between the maximum and minimum values in each window. This time-series, with an effective sampling frequency of 100 Hz, formed the envelope. It was filtered using a median filter with a window of 35 samples to remove impulsive noise, and then lowpass filtered with a 2^{nd}-order Butterworth filter with a cut-off frequency of 2 Hz, to obtain a smoothed envelope adjusted to the contour of breaths (Fig. 3, upper panels, yellow line).

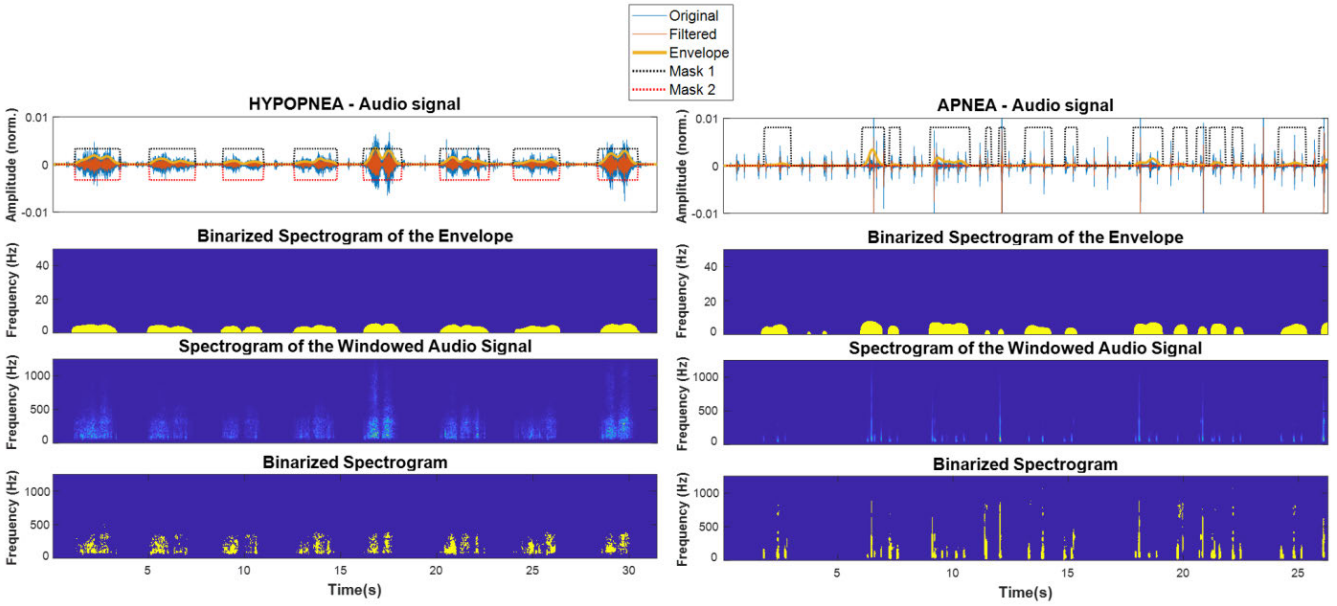

FIGURE 3. Depiction of the algorithm for classifying apneas/hypopneas, showing a case corresponding to hypopnea (left) and one corresponding to apnea with impulsive artifacts (right). From top to bottom: 1) audio signals (before filtering, in blue, and filtered in red) with their envelopes (thick yellow line, scaled ×10 in the apnea panel for easier visualization); 2) binarized spectrogram of the envelope; 3) spectrogram of the audio signal windowed with the first mask; and 4) binarization of this spectrogram. The binarized spectrogram of the envelope was used to select regions that could correspond to low-intensity breathing sounds or artifacts, and create a first mask (black dashed line in the first panel). Then, each window in this mask was checked in the time-frequency domain to remove short impulsive artifacts and build a final mask (red dashed line, first panel). This mask cannot be seen in the apnea case, because all the windows of the first mask contained artifacts and were thus discarded. The frequency limits of the audio signal spectrograms are shown up to 1250 Hz for easier visualization, since there is very little activity beyond this point.

To identify and delimit the potential regions of breathing activity, we computed a time-frequency representation (TFR) of the envelope. We chose the spectrogram as TFR, with time windows (Nwind) equal to 1% of the total segment length, 90% overlap (Noverl), and NFFT = 1024 points for the Fast Fourier Transform (FFT) calculation. This TFR was treated as an image, and binarized using a threshold equal to the 95th percentile of all the intensity values (Fig. 3, 2nd panels from the top). Then, the spots on the binarized image with a minimum duration of 25 ms were selected and a first mask was built (Fig. 3, black dashed line on top panels), indicating the times where some activity was present. However, this activity could be attributed to either LRS or artifacts. For this reason, we proceeded to a second stage where we analyzed those regions in more detail.

For each window in the mask, we calculated the spectrogram of the audio segment within it (Nwind = 0.01 s, Noverl = 90%, NFFT = 1024) (Fig. 3, 3rd panels from the top). Again, we binarized the resulting colormap using the 95th percentile of the intensity values of this short segment as the threshold (Fig. 3, bottom panels). As shown in the TFR of the windowed audio signal, breaths present a wide pattern, extending over longer time periods, while most impulsive artifacts show a pattern separated into several impulsive motifs that last less time. Accordingly, those motifs in the binarized TFR that were shorter than 20 ms were considered artifacts and removed. If isolated, this kind of impulsive noise would have been removed previously, but if several artifacts are grouped together their contributions are cumulative, producing a confounding effect on the envelope. For this reason, this second examination was necessary. Using this procedure, a second mask was built (Fig. 3, red dashed line on top panels), corresponding to the episodes that we ultimately recognized as breathing activity (LRS).

The final step consisted of examining this new mask to compute the duration of silence regions, i.e., the distances between breaths. If no LRS were identified (when the whole mask was equal to zero, because no breathing activity was detected), then the duration of the silence was the entire length of the SEv. The decision rule that we applied for classifying SEv is the following:

A SEv is labeled as apnea if one of these criteria are met:

1. There are more than two silences >6 s between LRS
2. There is a silence >10 s
3. There is a silence >6 s and <10 s, and the total duration of the SEv is <3 times the duration of this silence

All other silence events were classified as hypopnea.

2) COMPARISON WITH THE REFERENCE SYSTEM

Once the classification had been completed, we compared the labels that our automatic algorithm had assigned to each event with the actual apnea/hypopnea labels from ApneaLink. In this case, we used 'apnea' as a positive class and 'hypopnea' as a negative class. The number of TP, TN, FP, and FN was quantified. To distinguish these parameters from

those calculated when comparing SEv, we used the subindex "apn", e.g., TP_{apn}, TN_{apn}. We computed the sensitivity, specificity, and the positive and negative predicted values for this classification (Se_{apn}, Sp_{apn}, PPV_{apn}, NPV_{apn}), as well as the accuracy (Acc), defined as:

$$Acc(\%) = 100\frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (8)$$

E. STUDY OF NASAL VS ORAL BREATHING

To investigate the issue of nasal vs. oral breathing, and learn to distinguish them, we acquired an additional database (DB2). Experiments were conducted with 10 different participants: 5 men and 5 women, aged 21 to 58, who had not previously been diagnosed with sleep apnea. While lying on a bed or stretcher, either at home or in the lab, they were asked to breathe first through their noses (with their mouths closed), and then through their mouths with a nose clip preventing nasal breathing. Two one-minute signals were recorded with the smartphone microphone: one containing nasal breathing and the other oral breathing. The smartphone position over the thorax and the acquisition settings were the same as those described for DB1.

The signal processing was also the same: the signals were downsampled to 5 kHz, band-pass filtered between 70 and 2000 Hz, and the power-line noise was removed. After that, we characterized them using spectral and time-frequency representations. To work in the frequency domain, the FFT was calculated, and a linear envelope extracted, using windows of 15 Hz. The spectrogram was again used as TFR, with a window of 0.1 s, 90% overlap, and NFFT = 1024. The information extracted in this analysis was then used to further classify the SEv that we had previously labeled as 'hypopneas' into nasal or oral breathing. We also studied what these episodes looked like in the flow channel of the reference system.

III. RESULTS

A. SCREENING PERFORMANCE: AHI COMPARISON

As shown in Table 2, our database comprises 1269 AHEv, according to the reference system, of which 676 are apneas and 593 hypopneas. The smartphone detected 1246 SEv, applying the algorithm proposed in the previous section.

To evaluate the potential of smartphone audio signal acoustic analysis for screening OSA patients, once the SEv had been detected, we computed the predicted AHIs and compared them with the actual AHIs provided by ApneaLink. This comparison is presented in Fig. 4 and Table 2, which also shows, for each subject, the total duration of the analyzed signals, the number of apnea and hypopnea events in the reference system, and the total number of OSA-related events according to each method.

As illustrated in Fig. 4, the correlation between the AHIs of the two systems is excellent. The concordance correlation coefficient is very high: $\rho_c = 0.9923$. The AHIs predicted from the audio signals agree with the ApneaLink

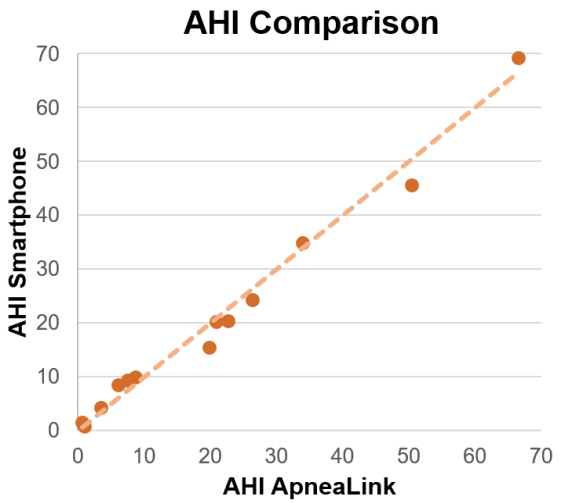

FIGURE 4. Comparison of the AHI (h^{-1}) predicted by the reference system (ApneaLink) and the smartphone audio signals. The concordance correlation coefficient between these variables is 0.9923. The orange dashed line represents the line y = x.

TABLE 2. Comparison of number of events and predicted AHI.

Subject	Duration (h:mm:ss)	Nº Apn Ref.	Nº Hpn Ref.	Nº Events Ref.	Nº Events Audio	AHI Ref. (h^{-1})	AHI Audio (h^{-1})
1	4:40:04	54	39	93	72	19.9	15.4
2	4:12:53	2	94	96	86	22.8	20.4
3	5:05:41	5	26	31	43	6.1	8.4
4	5:08:30	153	22	175	179	34	34.8
5	4:34:24	0	3	3	7	0.7	1.5
6	4:54:19	0	5	5	4	1	0.8
7	5:08:38	18	118	136	125	26.4	24.3
8	4:57:59	13	91	104	100	20.9	20.1
9	3:54:50	4	30	34	39	8.7	9.9
10	4:13:50	0	15	15	18	3.5	4.2
11	4:56:21	156	93	249	225	50.4	45.6
12	4:18:39	261	26	287	298	66.6	69.1
13	5:23:38	10	31	41	50	7.6	9.3
Mean / Total	4:43:48	676	593	1269	1246	20.7	20.3

Signal duration, number of apnea/hypopnea events and AHI predicted by the reference system (ApneaLink), and smartphone audio signals. Mean values of signal durations and AHIs, and total number of events can be found in the last row. Colors indicate OSA severity, as in Table I.

stratification of subjects into the 4 OSA severity levels: 3 normal (AHI < 5), 3 mild OSA ($5 \leq$ AHI < 15), 4 moderate OSA ($15 \leq$ AHI < 30), and 3 severe OSA (AHI ≥ 30). For these subjects, therefore, both the smartphone and ApneaLink would lead to the same diagnostic decision.

B. SILENCE EVENT COMPARISON

To investigate whether, apart from predicting similar AHIs, the events detected by the two systems (smartphone and ApneaLink) were comparable, we studied their agreement on an event-by-event basis.

The proposed automatic detector accurately identified the regions where silence or very low-intensity sound was present in all the full-night acoustic recordings. The automatically

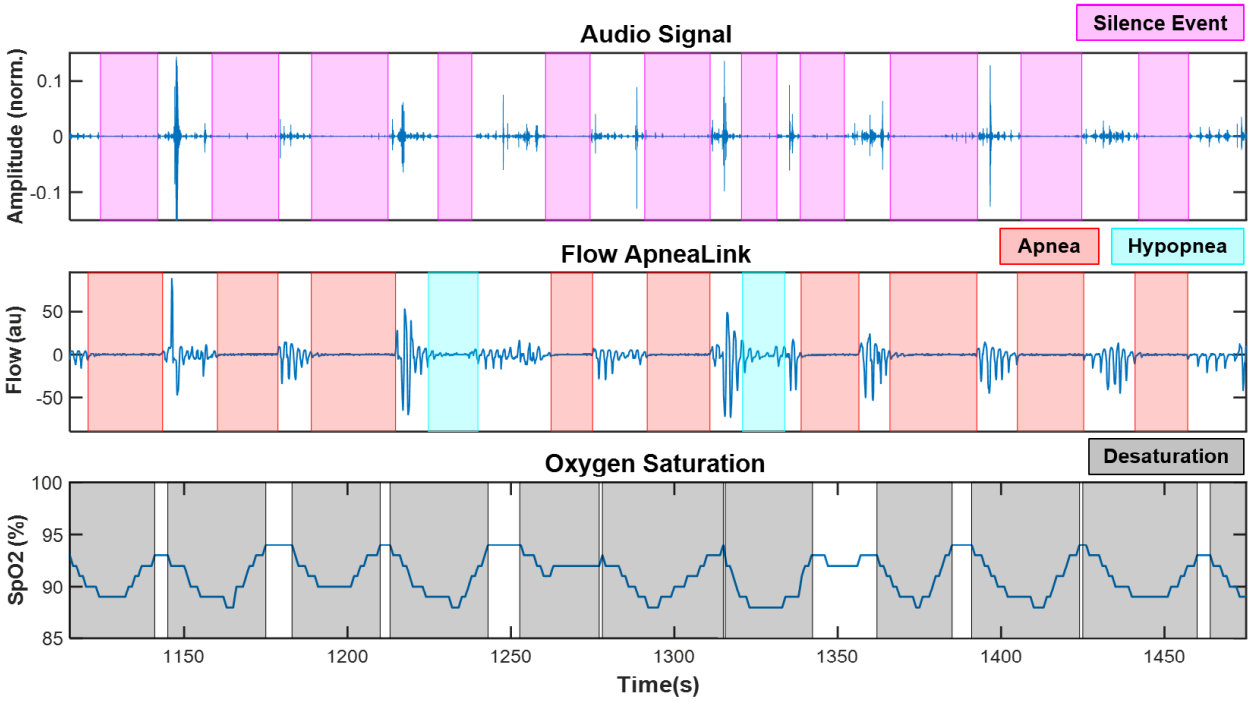

FIGURE 5. Example of a 6-minute segment showing the audio signal from the smartphone (upper panel), the nasal airflow channel (middle panel), and the pulse oximetry signal (bottom panel) from the reference device (ApneaLink). The silence events identified by our automatic detector are represented by the magenta boxes. They match the apnea (red boxes) and hypopnea (cyan boxes) events labeled in the flow channel. Desaturations resulting from AHEv are shown in grey boxes in the lower panel.

labeled SEv were consistent with visual inspection. Moreover, in general, these SEv correlated well with the AHEv of ApneaLink. An example is shown in Fig. 5, which shows a representative 6-min segment with many apneas and two hypopneas, comparing the airflow channel of the reference system and the smartphone audio signal. In this case, all the apneas and hypopneas match the SEv in the audio channel.

Table 3-A summarizes the results of the event-by-event comparison in terms of Se, PPV and percentage of apneas and hypopneas that were detected as SEv, for each patient and for the full database (total number of events). Table 3-B is the confusion matrix containing the total number of TP, FP, and FN. While the method demonstrates a high sensitivity to apneas, capturing almost 90%, and has a good PPV (82%), sensitivity to hypopneas is much lower: only 61% are detected as SEv. This affects the total sensitivity, which is 76% for the whole database, but in some subjects, particularly 1, 7 and 8, is quite low ($\leq$65%).

To investigate this issue, we examined what was happening in the acoustic signals in the segments corresponding to the FN of hypopneas (i.e., the hypopneas that our detector did not find). We discovered that in these regions there was no silence, as there were high-intensity sounds that clearly exceeded the threshold we had set in SampEn. In fact, during many of these hypopneas, snoring was present, as illustrated in Fig. 6. To quantify the recurrence of this phenomenon, we took the labels of snoring events provided by the reference system and checked how many hypopneas contained snores. Specifically, we considered 'hypopneas with snoring' to be only those which did not present a region >6 s without snores (otherwise, they could have been identified by our detector). The results are presented in the first two columns of Table 4-A, which displays, for each subject, the number of hypopneas with snoring vs. the total number of hypopneas that we did not detect. A total number of 119 episodes (more than half the FN of hypopneas) were associated with snoring, demonstrating that this is a prevalent anomaly.

We recalculated the event-by-event comparison excluding those hypopneas with snoring (Table 4). This was done to give us an idea of how our automatic algorithm performed in terms of detecting actual silence events. The total number of FN was greatly reduced with this procedure (from 296 to 177), and both the sensitivity to hypopneas and the global sensitivity improved greatly (%Hpn = 75%, Se = 84.3%). The algorithm therefore correctly identifies silence regions, but this is not sufficient to detect all the hypopneas, since many are accompanied by snoring. This is a clear limitation of the acoustic system that we must consider.

C. APNEA/HYPOPNEA CLASSIFICATION

As the acoustic signals from the proposed system can be very sensitive to low-intensity sounds, we went one step further by zooming in on SEv to try to classify them into hypopneas or apneas. We worked only with the 950 TP from Table 3, of which 601 (63.3%) corresponded to apneas in the reference system, and 349 (36.7%) to hypopneas. When we

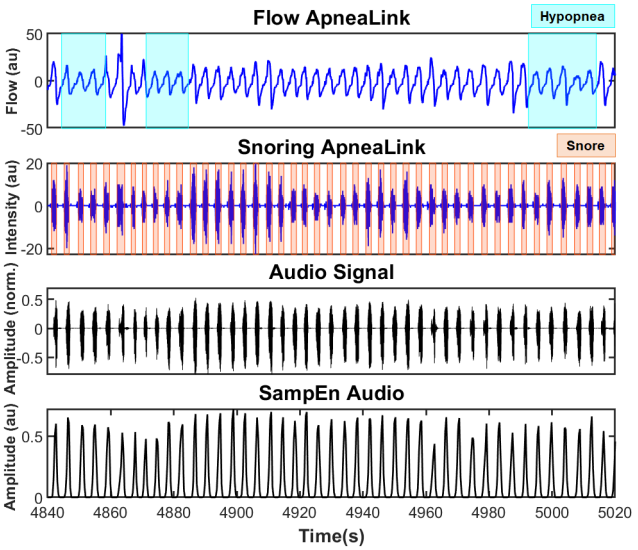

FIGURE 6. Segment showing that some hypopneas are accompanied by snoring. Hypopneas (cyan boxes) are indicated in the flow channel (upper panel), but these overlap with snores (orange boxes), as shown in the snoring channel of the reference system (2nd panel). The audio signal recorded with the smartphone microphone (3rd panel) is very similar, comprising high-amplitude snores. In these cases, the flow reduction does not produce a sound reduction, so there are no silence regions >6s in the SampEn envelope (lower panel). Therefore, these hypopneas with snoring cannot be detected as SEv.

TABLE 3. Event-by-event comparison: silence events vs. apnea/hypopnea events from the reference system.

A)

Subject	Se (%)	PPV (%)	%Apn	%Hpn
1	65.59	87.14	77.78	48.72
2	71.88	77.53	50.00	72.34
3	74.19	56.10	60.00	76.92
4	81.71	83.14	83.66	68.18
7	52.94	61.54	94.44	46.61
8	64.42	75.28	100	59.34
9	82.35	77.78	100	80.00
11	71.89	84.83	83.97	51.61
12	96.86	96.53	96.93	100
13	73.17	62.50	100	64.52
Total	**76.24**	**81.83**	**89.05**	**61.21**

B)

Audio \ ApneaLink	1	0
1	950	211
0	296	-

Event-by-event comparison of the SEv automatically detected in the smartphone audio signals and the AHEv from the reference system (ApneaLink). Table III-A shows the values of sensitivity, positive predictive value, and percentage of apneas and hypopneas detected for each patient. Table III-B contains the total number of TP (1 in Audio and ApneaLink), FP (1 in Audio, 0 in ApneaLink), and FN (0 in Audio, 1 in ApneaLink).

found an acoustic activity that was consistent with breathing, this SEv was automatically classified as 'hypopnea', whereas when this respiratory activity was absent, it was labeled as 'apnea'. After that, the labels were compared to those from the reference system. The results can be seen in Table 5.

The total accuracy is 70%, which is rather low. If we built a basal model simply by assigning to all the events the label

TABLE 4. Event-by-event comparison without 'Hypopneas with Snoring'.

A)

Subject	Nº Hpn + Snor[a]	Nº FN Hpn[b]	Se (%)	PPV (%)	%Apn	%Hpn
1	13	20	76.25	87.14	77.78	67.86
2	9	26	79.31	77.53	50.00	79.07
3	0	6	74.19	56.10	60.00	76.92
4	4	7	83.63	83.14	83.66	75.00
7	42	63	76.70	61.54	94.44	67.90
8	23	37	82.72	75.28	100	77.14
9	3	6	90.32	77.78	100	85.71
11	24	45	79.56	84.83	83.97	66.67
12	0	0	96.86	96.53	96.93	100
13	1	11	75.00	62.50	100	76.92
Total	119	221	**84.29**	**81.83**	**89.05**	**75.00**

B)

Audio \ ApneaLink	1	0
1	950	211
0	177	-

Quantification of hypopneas with snoring and event-by-event comparison after excluding these events. Table IV-A shows the values of Se, PPV, %Apn, and %Hpn for each patient. Table IV-B contains the total number of TP, FP, and FN.

[a]Nº Hpn + Snor: number of hypopneas with snoring

[b]Nº FN Hpn: number of FN corresponding to hypopneas, i.e., hypopneas from the reference system that were not detected as SEv. There was a total of 296 FN (Table III-B), of which 221 were hypopneas and 75 apneas.

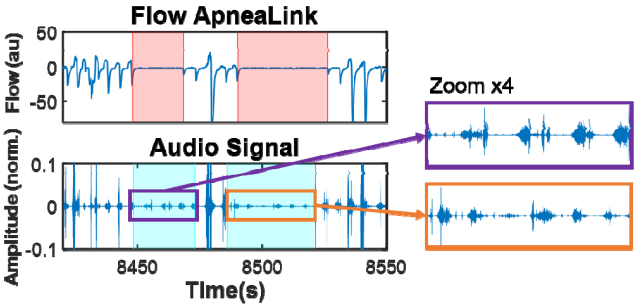

FIGURE 7. Example of events that are labeled as 'apnea' in the reference system (red boxes) but show breathing activity in the audio signal, thus being classified by our automatic algorithm as 'hypopneas' (cyan boxes). Although there is an absence of nasal airflow during these events, we can still observe clear low-intensity breathing sounds in the acoustic signal from the smartphone. The figures on the right are an expansion of these events, zooming amplitude ×4. This phenomenon could correspond to oral breathing omitted by the nasal cannula.

of the class with the highest *a priori* probability (in this case, apnea is the most prevalent class), this would already have an accuracy of 63.3%. There is, therefore, little improvement over the basal model.

If we look at the numbers in Table 5, the worst values are found for Sp (67%) and NPV (58%), which are related mainly to hypopnea detection errors. Note that specificity to apneas is equivalent to sensitivity to hypopneas.

As reported in Table 5-B, there is a high number of FN_{apn}, i.e., events classified as hypopnea that the reference system labeled as apnea. Patient 4 accounts for the majority of these errors, which drastically reduce the total accuracy, since he is a severe subject, with many apneas, most of which we classified as hypopneas ($Se_{Subj4} = 39\%$). However, exploring the audio signals during these events, we noticed that

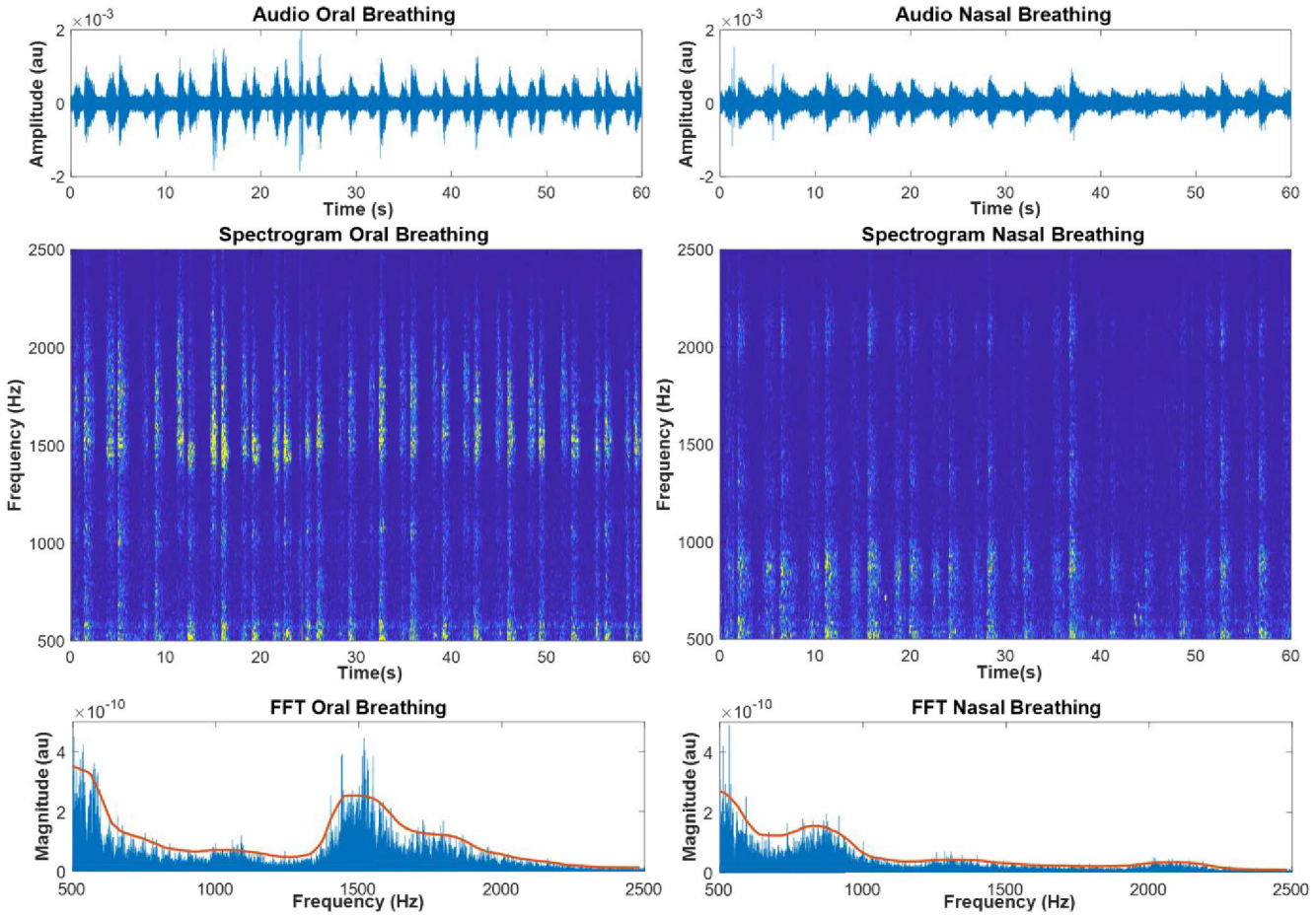

FIGURE 8. Time, time-frequency and spectral representations of 1-minute segments of oral (left) and nasal (right) breathing from one of the study participants (Subject 1 from DB2). The top panels represent the signals in the time domain, followed by spectrograms (middle panels), and finally the FFTs with their envelopes (red) (lower panel). Oral breathing presents a characteristic peak at higher frequencies (in this case, around 1.5 kHz) that is not seen in nasal breathing.

we could clearly see breathing within most events labeled as apnea in the reference system but classified as hypopnea by our detector. An example of this is in Fig. 7. We suspected that these events might contain oral breathing that had not been recorded by the nasal cannula. We therefore proceeded to characterize the oral and nasal breathing acoustic signals using a new database, to make a detailed analysis of what was happening in these mislabeled events.

D. NASAL VS ORAL BREATHING

We used the signals from DB2 to characterize oral and nasal breathing and explore whether we can distinguish them easily. To do this, we explored the signals in the time, frequency, and time-frequency domains (Fig. 8). As a result, we found that oral breathing has a very specific spectral pattern, clearly distinct from that of nasal breathing.

As can be seen in the spectrograms and FFT in Fig. 8, most of the power of nasal breathing is concentrated in the low-frequency band, decreasing sharply after 1 kHz. In contrast, the oral breathing spectrum presents a characteristic prominent peak at around 1500 Hz. Moreover, a similar pattern is present in all the cases we examined. The exact location of the oral breathing peak shifts from one individual to another, but is always found between 950 Hz and 2 kHz. The spectral pattern of nasal breathing is also very repetitive between individuals, and never shows this kind of peak within this frequency range. This enables us to discriminate oral and nasal breathing.

To verify these assumptions, we looked for the maximum peak in the FFT envelope in the mentioned frequency band (950-2000 Hz). The frequency where these peaks are located (fpeak), and their heights (Peak Height %) are shown in Table 6. The peak height was normalized to the maximum height of the envelope in the frequency band 500-2000 Hz, to avoid the variable effect of basal noise, which has a high magnitude and is concentrated at low frequencies. These features were calculated for each 10-second window from the 1-minute recordings from DB2, so, for each patient, we had 6 windows for oral breathing and 6 for nasal breathing. The results in Table 6 are summarized as the mean and standard deviation of each parameter for each subject, and for the total database.

TABLE 5. Results of Apnea/hypopnea classification. direct comparison with the reference system.

A)

Subject	Se_{apn} (%)	Sp_{apn} (%)	PPV_{apn} (%)	NPV_{apn} (%)	Acc (%)
1	83.33	73.68	87.50	66.67	80.33
2	100	69.12	4.55	100	69.57
3	33.33	60.00	11.11	85.71	56.52
4	39.06	46.67	86.21	8.24	39.86
7	64.71	85.45	57.89	88.68	80.56
8	84.62	85.19	57.89	95.83	85.07
9	75.00	54.17	21.43	92.86	57.14
11	79.39	52.08	81.89	48.08	72.07
12	83.00	36.00	92.92	17.31	78.78
13	50.00	70.00	45.45	73.68	63.33
Total	**71.59**	**67.24**	**79.08**	**57.78**	**70.00**

B)

Audio \ ApneaLink	Apn	Hpn
Apn	431	114
Hpn	171	234

Comparison of the apnea/hypopnea labels assigned by the automatic algorithm and the apnea/hypopnea labels from the reference system. Table V-A shows the values of sensitivity, specificity, positive predictive value, and negative predictive value for each patient, with 'apnea' being the positive class. Table V-B shows the total number of TP_{apn} ('apnea' from both the audio and ApneaLink), TN_{apn} ('hypopnea' from both the audio and ApneaLink), FP_{apn} ('apnea' from the audio, but 'hypopnea' from ApneaLink) and FN_{apn} ('hypopnea' from the audio, but 'apnea' from ApneaLink).

TABLE 6. Spectral Features For characterizing nasal & oral breathing.

	ORAL BREATHING		NASAL BREATHING	
Subject DB2	fpeak (Hz)	Peak Height (%)	fpeak (Hz)	Peak Height (%)
1_2	1510±40	91±14	1250±250	17±10
2_2	1510±40	87±13	1320±70	18±3
3_2	1500±15	87±15	1570±220	11±12
4_2	990±30	100±0	1570±300	6±6
5_2	1570±70	100±0	1490±240	15±3
6_2	1200±60	100±0	980±590	23±16
7_2	1370±240	70±10	1370±210	9±5
8_2	1470±25	93±8	1320±50	27±2
9_2	1210±45	100±0	1470±320	25±5
10_2	1350±45	88±14	1220±100	45±7
Total	**1370±190**	**92±12**	**1350±300**	**20±13**

Peak frequency and height of this peak (in % of the maximum value in the bandwidth 500 – 2000 Hz) for each subject in nasal and oral breathing. These parameters were calculated from the PSD of 10-s windows. They are expressed as mean ± standard deviation.

Table 6 confirms that oral breathing always presents a high peak in the explored band. Its peak height is either the maximum value in this frequency band or at least 70% of this value (>85% in all cases except subject 7_2). However, this peak does not exist in nasal breathing, or it is much more attenuated, rarely reaching values >30%, and never >50%. The peak location (fpeak) is subject-specific and varies between subjects, especially in subject 4_2, whose fpeak is around 1 kHz. In nasal breathing there is no similar peak, only residual activity (Fig. 8), so fpeak can be almost any value and there is greater variability between the different windows of the same subject.

TABLE 7. Classification of events labeled as 'hypopneas' by the automatic detector into nasal or oral breathing.

ApneaLink Label	APNEAS (FN_{apn})		HYPOPNEAS (TN_{apn})	
Breathing / Subject	ORAL	NASAL	ORAL	NASAL
1	3	4	1	13
2	0	0	0	47
3	0	2	0	12
4	70	8	3	4
7	1	5	1	46
8	0	2	0	46
9	0	1	0	13
11	2	25	0	25
12	36	7	5	4
13	0	5	2	12
Total	**112**	**59**	**12**	**222**

Number of events classified as nasal or oral breathing for each subject, from the total number of SEv that we classified as 'hypopneas', separating FN_{apn} (those labeled as 'apnea' in ApneaLink) from TN_{apn} (those labeled as 'hypopnea' in ApneaLink). Most FN_{apn} correspond to oral breathing.

Since the differences seemed so evident, we decided to use just one feature, the normalized peak height in this frequency band, to discriminate between nasal and oral breathing. According to the values obtained in Table 6, we set the threshold for the peak height as 60% of the maximum. To distinguish oral and nasal breathing, we computed the envelope of the FFT and looked for the maximum peak in the 950-2000 Hz band. If that peak had a height reaching the threshold, then it was labeled as oral breathing, otherwise it was considered nasal breathing. In this training database, the proposed method had 100% classification accuracy.

We took all the SEv that our algorithm had classified as 'hypopneas' (i.e., the 171 FN_{apn} and 234 TN_{apn} from Table 5-B) and further classified these as oral or nasal breathing. The results are shown in Table 7, which also separates these events according to the reference system label ('apnea' corresponds to FN_{apn} and 'hypopnea' to TN_{apn}). While only a few TN_{apn} present oral breathing, many of the FN_{apn} are explained by this, especially in subjects 4 and 12. This supports the hypothesis that these events correspond to oral breathing that cannot be recorded with a nasal sensor, so the reference system incorrectly labels them as 'apnea', despite them being hypopneas as there is, indeed, breathing activity.

To evaluate the performance of our algorithm for LRS detection and apnea/hypopnea classification, we recalculated the comparison from Table 5, changing the 112 reference system labels corresponding to oral breathing from 'apnea' to 'hypopnea'. The results are given in Table 8. Se_{apn} and NPV_{apn} show a large increase, as we now detect most of the apneas, and when our algorithm labels an event as 'hypopnea', it is generally correct. The total accuracy is now 81.8%, clearly better than the basal model. The Sp_{apn}, which indicates our ability to detect hypopneas, and the PPV_{apn} have good but slightly lower values. This is because, for some subjects and in some parts of the night, there may be some artifacts, the SNR is too low, or the magnitude of the signal is too low, so it is not possible to discriminate

TABLE 8. Results of Apnea/hypopnea classification after correcting the oral breathing labels.

A)

Subject	Se_{apn} (%)	Sp_{apn} (%)	PPV_{apn} (%)	NPV_{apn} (%)	Acc (%)
1	89.74	77.27	87.50	80.95	85.25
2	100	69.12	4.55	100	69.57
3	33.33	60.00	11.11	85.71	56.52
4	86.21	90.59	86.21	90.59	88.81
7	68.75	85.71	57.89	90.57	81.94
8	84.62	85.19	57.89	95.83	85.07
9	75.00	54.17	21.43	92.86	57.14
11	80.62	54.00	81.89	51.92	73.18
12	96.77	73.77	92.92	86.54	91.73
13	50.00	70.00	45.45	73.68	63.33
Total	**87.96**	**75.22**	**79.08**	**85.43**	**81.79**

B)

Audio \ ApneaLink	Apn	Hpn
Apn	431	114
Hpn	59	346

Comparison of the apnea/hypopnea labels assigned by the automatic algorithm and the apnea/hypopnea labels from the reference system, after changing the 'apnea' labels corresponding to oral breathing to 'hypopnea'. Table VIII-A shows the values of sensitivity, specificity, positive predictive value, and negative predictive value for each patient, with 'apnea' being the positive class. Table VIII-B contains the new values for TP_{apn} ('apnea' from both the audio and ApneaLink), TN_{apn} ('hypopnea' from both the audio and ApneaLink), FP_{apn} ('apnea' from the audio, but 'hypopnea' from ApneaLink), and FN_{apn} ('hypopnea' from the audio, but 'apnea' from ApneaLink).

respiratory activity and these SEv are incorrectly classified as apneas. This is especially critical in subject 2, who has many hypopneas but ∼30% are misclassified, affecting the global results. Nonetheless, this issue is uncommon, and the global results demonstrate that the developed algorithm accurately classifies apneas and hypopneas, being especially sensitive to apneas (88%) but also to hypopneas (75%).

IV. DISCUSSION

A. ACOUSTIC ANALYSIS FOR MONITORING BREATHING ACTIVITY

In this work, we have presented an algorithm based on the SampEn of audio signals to identify breathing cycles and silence events corresponding to either apneas or hypopneas. We chose to use SampEn because it is a robust measurement suitable for noisy and complex physiological signals [32]. In our case, this is particularly important because the data come from acoustic signals acquired with an ambient microphone, but this should pick up very low intensity breathing sounds, so the SNR may be very low. One of the main advantages of SampEn is its resistance to impulsive noise [36]. This property is very useful for correctly identifying SEv, as they are usually contaminated by small impulsive artifacts from cardiac activity, swallowing, clicking, and other sounds from the mouth and the upper airways [31].

Previous sleep apnea studies used entropy measurements from acoustic signals recorded with contact microphones to extract different features and build classification models [37], [38]. Our proposal also involves an entropy analysis of acoustic signals, but acquired by a smartphone microphone. Our algorithm uses SampEn as an envelope estimator to directly track respiratory events. Although our purpose was only to detect SEv, the same approach could easily be adapted to identify each respiratory cycle, corresponding to a single breath or snore. This would allow continuous monitoring of the respiratory rate overnight and facilitate a snoring analysis. This approach would have multiple applications, not only providing more information for OSA management, but also for tracking sleep and assessing other respiratory diseases.

B. OSA SCREENING WITH AUDIO SIGNALS FROM A SMARTPHONE

The SEv detected by our automatic algorithm in audio signals were used to predict the AHI of each subject. These values were compared with the AHIs provided by the reference system and used to evaluate the subjects' condition. We assessed whether they had OSA or not, and their level of severity. The concordance between the AHIs predicted by the two systems (smartphone and ApneaLink) was excellent (Fig. 4), demonstrating the promising performance of the proposed method.

Our sample (13 subjects) included males and females of different ages. Of these, 3 females proved to be healthy asymptomatic individuals, and 10 (2 females and 8 males) had OSA, covering all severity types (mild, moderate, and severe). This diversity allowed us to evaluate the potential of our system across a wide range of sleep apnea conditions. OSA is known to be more prevalent in men and the elderly population; our 3 healthy subjects are young females, while 2 of the 3 mild patients are much younger men than the moderate and severe patients. Nevertheless, the sample size is insufficient to extract general age- or gender-related conclusions. This issue is beyond the scope of this work, as it is not intended to be a population study, but rather a proof-of-concept of the feasibility of using the smartphone system for OSA screening.

The number of events found by ApneaLink and smartphone was similar for all subjects (Table 2), regardless of whether the patient had mainly apneas (e.g., the most severe subjects 4, 11, 12) or hypopneas (e.g., subjects 2, 7, 8). This suggests that SEv from the audio signals may correspond to both apneas and hypopneas, and that we are therefore capturing both types of episodes.

A very positive result is that all subjects were correctly diagnosed: all the healthy subjects were classified as healthy, and all the OSA patients were identified. Furthermore, all patients were correctly stratified according to their level of severity (mild, moderate, or severe). Therefore, in our database, the smartphone performs equivalently to ApneaLink in terms of detecting and classifying OSA patients. Moreover, there is a clear separation between the predicted AHIs of mild and moderate OSA subjects (Fig. 4),

allowing easy identification of the patients at the highest risk and with an urgent need for treatment. Distinguishing between healthy and mild OSA subjects is usually more challenging, but in our case, it does not pose a problem and the classical threshold of AHI = $5h^{-1}$ is still valid.

C. PULSE OXIMETRY, A NECESSARY GUIDE

The algorithm that we developed for SEv detection required the pulse oximeter information to select the segments to be analyzed. This helped to reduce the computational cost, and particularly the number of false alarms. Without this guide, we could overestimate the severity of a patient. This may appear to be a limitation of the acoustic signal, as this information alone cannot provide such an accurate analysis. However, there is a similar situation in clinical practice. OSA-related events are usually labeled on the airflow channel, but, while apneas can be directly annotated when there is a flow reduction $\geq$90%, hypopneas must be followed by desaturation or arousal. Therefore, SpO_2 also guides the annotation of events in clinical practice.

We mimicked the clinical procedure, by taking advantage of SpO_2 to help us detect SEv in the audio signals. The main problem of this approach is that we may miss apneas that are not followed by desaturation. Nevertheless, this rarely happens, since apneas are usually associated with a decrease in oxyhemoglobin saturation [39], [40]. In fact, obstructive apneas are supposed to lead to longer, greater in area and deeper desaturations than those caused by hypopneas [41]. For this reason, the events that we might miss would primarily correspond to short apneas associated with less serious physiological effects, as they do not produce hypoxia.

The SpO_2 information that we used came from the pulse oximeter of ApneaLink, so the smartphone system was not autonomous. We chose to use the desaturations from ApneaLink because we wanted to examine the same regions in both systems. As we aimed to compare the information provided by the audio and nasal airflow signals, we needed to avoid any effects introduced by different SpO_2 sensors. If we had connected an external pulse oximeter to the smartphone, we would not have been able to ascertain the origin of the discrepancies in the results. Some of the errors would come from our algorithms and audio signals, but others would come only from the differences in the dynamics of the two pulse oximeters and separating out their contributions would be difficult.

In fact, a wired or wireless fingertip pulse oximeter could easily be integrated into the smartphone system to make it autonomous, as we did in previous pilot studies [30], [31]. Then, the same algorithm for SEv could be applied, guided by the desaturations detected in the SpO_2 signal from this new sensor. We will do this in future work, but this new pulse oximeter must be accurately calibrated. After taking these steps, integrating these and perhaps further sensors into the smartphone platform could lead to a more robust mHealth prototype.

D. AUTOMATIC APNEA AND HYPOPNEA DETECTORS

We have developed automatic algorithms for detecting SEv in audio signals and classifying them as apneas or hypopneas. OSA is usually diagnosed after manual inspection of the full-night PSG signals, which is a time-consuming and exhausting task, even for the most experienced specialists. Automatic screening tools, such as that in the proposed approach, using audio signals from smartphones, could help reduce human error and workload, support clinical decision-making, and improve diagnostic efficiency and quality.

To test the performance of our algorithms, we performed an event-by-event comparison against ApneaLink. The results confirmed that SEv have a physiological sense, since most correspond to AHEv from the reference system (PPV = 82%, Table 3). The sensitivity of our system to ApneaLink events was 76%. It captured 89% of the apneas, but only 61% of the hypopneas. This is explained by the fact that many hypopneas contained snoring, as we will discuss in a later section.

There was a total of 296 FN (Table 3-B), of which 221 corresponded to hypopneas and 75 to apneas. A higher sensitivity to apneas than hypopneas was expected, because the definition of hypopnea is more complex and remains controversial [42]. Indeed, hypopneas can be any flow reduction between 30% and 90%, and it is unlikely that this entire range of events exhibits the same acoustic characteristics.

More than half the missing hypopneas (119/221) were accompanied by snoring, but the remainder also had breathing activity beyond the threshold that we had set in SampEn. A considerable number of the FN corresponding to apneas (32/75) contained high-intensity respiratory sounds compatible with oral breathing, while the remaining errors could be attributed to sound artifacts. The SEv detector does, therefore, accurately determine silence regions, but each signal has its own characteristics and resolution, and a flow reduction does not always imply a sound reduction. This explains the discrepancies when breathing or snoring sounds are present, even though there is a significant reduction in the nasal flow channel.

On the other hand, the FP errors have various explanations. Firstly, we are using events longer than 6 s, but ApneaLink only detects events longer than 10 s. Also, as we explained previously, our algorithm is very sensitive to low-intensity sounds, so, when there are long hypopneas in airflow, it tends to detect multiple short SEv instead. This behavior can produce FP if some of these short SEv do not overlap the event in the reference system. Finally, in segments of the night where the signal magnitude is too low, it can be complicated to detect actual silence regions, so the algorithm can make errors. Despite these facts, the total PPV is 82%, so we can be confident that the vast majority of our automatically detected SEv are really AHEv.

Regarding the classification of apneas/hypopneas, the main discrepancies between our algorithm and the reference device

arose from the inability of the nasal cannula to capture oral flow, as we will discuss below. Apart from this issue, the results are encouraging, since, after correcting these labels, the algorithm achieves a classification accuracy of 82% (Table 8). The Se_{apn} is 88%, which means that our method correctly classifies most apneas as 'apneas'. The NPV_{apn} is 85%, which means that we are usually right when we label an event as 'hypopnea'. The slightly lower values of $Sp_{apn} = 75\%$ and $PPV_{apn} = 79\%$ suggest that, in some cases, we are still unable to detect hypopneas. This is probably an issue of sensor resolution, because, at certain times during the night, the magnitude of the signal is too low, so it is impossible, even visually inspecting the data after filtering, to detect breathing activity. This kind of errors were uncommon, except in subject 2, who is an outlier but has many hypopneas, of which about 30% are misclassified as apneas (Table 8).

If we look at Table 8, we can see that the worst results are those for the mild subjects (3, 9, 13), where the accuracy is <65%. There are various reasons for this: 1) these subjects have a low number of events, so a few errors are sufficient to significantly impact the classification accuracy estimators; 2) the events from these subjects, mostly hypopneas, are not as clear as in more severe subjects, i.e., there is not as much flow reduction or sound reduction, and they are usually located in noisy regions, so they are more difficult to detect; and 3) these subjects do not usually snore, or snore less than more severe patients, so the SNR is worse due to the lower amplitude of breathing sounds. Nonetheless, the errors in this population are not as critical, because these patients are at a lower risk, as indicated by their AHIs. Due to their low number of events, their contribution to the general results in the complete database is almost imperceptible.

Overall, the proposed algorithm demonstrates a high level of accuracy, being very sensitive to apneas, the most critical events, as well as to hypopneas. This solution could easily be implemented using portable devices or smartphone apps, as it is a quick and simple tool for classifying apneas and hypopneas. Although OSA diagnosis is primarily based on just the AHI, distinguishing between apneas and hypopneas has clinical value. Incorporating this information would help the physician to better understand the patient's pathophysiology and prognosis and make a therapeutic decision.

E. SNORING IS PRESENT DURING SOME HYPOPNEAS

When we directly compared the SEv from the audio signals with the AHEv from the reference system, we noticed that our algorithm missed many hypopneas. We realized that, in fact, most of these hypopneas were accompanied by snoring. This phenomenon is already considered in the AASM guidelines, which mention 'obstructive hypopneas with snoring' as a subtype of hypopneas. These descriptions are based on the snoring channel derived from a nasal pressure transducer, but this issue has not been studied directly in acoustic signals.

One might expect both apneas and hypopneas to produce a significant sound reduction. However, during hypopneas there is still a certain acoustic activity, making it possible to differentiate them from apneas. Indeed, in some cases, this acoustic activity has such a high magnitude that the hypopnea can no longer be labeled a SEv. Among these events, we find the hypopneas accompanied by snoring, as we saw clearly in both audio signals and the snoring channel of ApneaLink (Fig. 6). Therefore, during snoring hypopneas, the flow reduction is not reflected by a reduction in sound.

We recently observed the behavior of hypopneas with snoring in audio signals, for the first time, in a pilot study [31], but here we confirm this with a larger database. This kind of hypopnea justifies the relatively low sensitivity (76%) that we obtained in the direct comparison (Table 3), which noticeably improves (84%) after excluding these events (Table 4). Although this fact points to a limitation of our mHealth prototype and acoustic analysis for detecting hypopnea, it also evidences that the proposed algorithm accurately detects SEv. Nonetheless, SEv are not sufficient to encompass all hypopneas.

The existence of snoring during some hypopneas means that, although there is an airflow reduction, the passage of this remaining flow through the partially blocked upper-airway produces a vibration in the anatomical structures, leading to these irregular high-intensity sounds. However, the exact mechanisms are unknown. Moreover, it remains an open question as to why, in some hypopneas, the flow reduction leads to reduced sound, while in others it produces snoring. Further investigation is needed to clarify this issue.

Currently, our algorithm cannot be used to detect hypopneas with snoring, since it was specifically designed to detect silence events. This issue should be further explored using other types of analysis to assess whether it is possible to distinguish regular snoring from that which occurs during hypopneas. Despite this limitation, our automatic SEv detector has considerable potential for OSA monitoring, as it accurately identifies apneas, the most acute events, and 'silence hypopneas' (those where no snoring is present and which can be detected as SEv). 'Silence hypopneas' are the most similar to apneas, and therefore these are likely to be the most critical in terms of physiological effects. Nevertheless, further research is needed to test these hypotheses.

F. AUDIO SIGNALS CAN DETECT ORAL BREATHING

We have characterized the acoustic signals corresponding to oral and nasal breathing. Their spectral content enabled us to easily distinguish the breathing route. Furthermore, the spectral pattern of oral breathing was shared by most of the events

where we saw breathing activity in the acoustic channel, despite being labeled as 'apnea' by the reference system (FN_{apn}, Table 7). In contrast, almost all the events labeled as 'hypopnea' both by the smartphone and ApneaLink (TN_{apn}, Table 7) had the characteristics of nasal breathing. The few remaining TN_{apn} (12 out of 234) could correspond to oronasal breathing. This evidence supports our hypothesis that many discrepancies between the smartphone and the reference system in terms of FN_{apn} can be explained by the inability of the nasal cannula to capture oral breathing.

The official AASM definition of apneas specifies that these events are reductions in ''oronasal airflow'' [6]. For this reason, we decided to change the labels of the events in the reference system that corresponded to oral breathing from 'apnea' to 'hypopnea', to better test the performance of our algorithm. The results were greatly improved, as reflected by the increase in the classification accuracy from 70 to 82%.

In this work, we used oral breathing information to explain the differences in the apnea/hypopnea classification between our automatic algorithm from audio signals and the reference system. However, the knowledge acquired on nasal and oral breathing patterns could have many more applications. For example, we could easily adapt the proposed approach to monitor, overnight, exactly when and how much the patient breathes through the nose or mouth. This could be of great clinical value because, as discussed in previous work, excessive breathing through the mouth can worsen the effects of OSA [11], but may also indicate nasal congestion, allergies, or be related to other illnesses [43].

Most of the oral breathing episodes in the nocturnal database came from subjects 4 and 12, who have severe OSA. This is in line with previous work associating increased oral breathing with OSA [10]. Since breathing through the mouth can be a risk factor for OSA and worsen the condition [11], even reducing the effect of therapeutic approaches [12], [13]; it would be useful to monitor the breathing route in patients with OSA over several nights. This could be done easily, in a non-invasive way, using the smartphone.

One of the aims of this study was to compare the information provided by acoustic signals from the smartphone with the airflow signals from the reference device. Oral breathing has proved to be a limitation for systems based on nasal airflow. We found a significant number of events where the audio signal clearly showed breathing activity, while nothing could be seen in the flow channel, and thus the reference system labeled these events as apneas (Fig. 7). This could lead to apneas and hypopneas being confused, but also to overestimation of patients' severity, if they breathe through their mouths a lot. This highlights the importance of having sensors able to record oral breathing so that we can provide an accurate OSA diagnosis. Sensors that satisfy this condition include, as shown in this work, high-quality microphones, but also, for instance, the oronasal thermistors that are usually used in PSG.

A similar discussion was presented in a previous work comparing OSA-related events scored with a nasal pressure transducer only, or including an oronasal thermal sensor [8]. The authors concluded that the diagnostic accuracy was affected by the absence of the oronasal thermal sensor, but that any impact was likely to be small. Our results suggest that, while it is true that the differences in AHI would not be critical, being able to record oral breathing is essential for distinguishing apneas and hypopneas.

G. SMARTPHONES AS SLEEP MONITORING TOOLS: ADVANTAGES, LIMITATIONS, AND FUTURE WORK

We have proposed a method for detecting apneas and hypopneas using acoustic signals recorded with a smartphone, and we have demonstrated its potential for screening and monitoring patients with OSA. However, this study is not free of limitations, which should be addressed by future research.

Most work on novel tools for detecting OSA compare their results with PSG, because it is the gold-standard technique in OSA diagnosis. In this study, as we propose a smartphone system whose final application would be the home-screening of OSA patients, we compared it with another portable device designed to be used at home. Future work could include a comparison between our system and PSG, too. Nevertheless, we would expect similar or better results, since the PSG tests would be performed in a more-controlled and less-variable environment than a patient's home, and under the supervision of qualified practitioners.

An advantage of the smartphone system is that it uses unobtrusive sensors, reducing the number of wires, and can be much more comfortable for patients. According to the feedback we received, 10 of the 13 subjects reported no discomfort or problems sleeping. The others felt that the inconvenience stemmed from sleeping with the two devices attached simultaneously to their chest, rather than because of the smartphone. On the other hand, all 13 participants confirmed that it was easy to activate the acquisition, as they simply had to press a button.

Possible next steps include integrating other built-in or external sensors to complement the information provided by the audio channel. In particular, we want to add a wired fingertip pulse oximeter to provide oxygen desaturation information and make the smartphone system autonomous.

Despite the limitations mentioned here, the acoustic analysis of smartphone signals has shown to be an automatic, straightforward and effective approach for identifying and stratifying OSA patients. The availability and low-cost of smartphones, linked with their powerful sensors, make them perfect mHealth devices for performing simple non-invasive diagnostic tests at home. In this way, they can facilitate the diagnostic procedure and contribute to reducing the huge costs and workload that OSA represents for healthcare systems.

The proposed system and other emerging mHealth devices are not necessarily aimed at substituting the gold-standard

technique, PSG. In fact, they should rather complement PSG, as screening tools for reaching many more patients and helping the early detection of people at risk of OSA or other sleep disorders. In addition, since smartphones are widely available devices, our approach could be used to monitor a patient's condition over several nights and study the variability in the AHI, something that is currently not possible due to the high resource cost. In future work, we will monitor OSA patients for several consecutive nights and/or, alternatively, for several nights distributed across different months, as a long-term follow-up. The smartphone will be very valuable for this application, which will allow continuous monitoring of the patient's condition and the effect of therapeutic interventions, thus optimizing the entire clinical process.

V. CONCLUSION

In this work, we have presented a smartphone-based system for screening and monitoring OSA patients at home using acoustic signal analysis. We achieved our first objective by developing an automatic algorithm for detecting silence events, which correlated well with the reference system labels for apneas and hypopneas.

Our second objective was to test the performance of the smartphone system for OSA detection. In our database, the AHIs predicted from the audio signals had a very high degree of concordance with those from the commercial reference device, and all the patients were correctly diagnosed and stratified into severity levels.

Finally, the third main objective was to compare the information provided by the audio signals from the smartphone microphone and the airflow channel of the reference device. Although the sound reductions in the audio signals could be used to identify apneas and some hypopneas, we found that certain hypopneas are accompanied by snoring, and thus cannot be detected as silence events. This challenge for the acoustic analysis must be considered in the development of future OSA diagnosis solutions.

Additionally, we provided the first acoustic characterization of oral vs. nasal breathing. Thanks to this, a limitation of the nasal airflow sensors was confirmed: they miss oral airflow, so they may mislabel episodes of oral breathing as apneas. This underlines the importance of having sensors able to record both nasal and oral airflow to accurately evaluate OSA.

This novel knowledge of oral breathing and hypopneas with snoring may be of great interest for clinical practice, to better understand the pathophysiological mechanisms of the events associated with OSA, but also the potential limitations of current and novel diagnostic equipment. An automatic analysis of acoustic signals from smartphones can be used as a powerful, non-invasive tool for early detection and monitoring of OSA patients at home, thus helping to reduce the high impact of this disease.

REFERENCES

[1] P. Escorrou, *Sleep Apnea: Implications in Cardiovascular and Cerebrovascular Disease*, vol. 20, T. D. Bradley and J. S. Floras, Eds., 2nd ed. London, U.K.: Informa Healthcare, 2011, p. 213.

[2] A. Yoshihisa and Y. Takeishi, "Sleep disordered breathing and cardiovascular diseases," *J. Atheroscler. Thromb.*, vol. 26, no. 4, pp. 315–327, Apr. 2019.

[3] D. J. Durgan and R. M. Bryan, Jr., "Cerebrovascular consequences of obstructive sleep apnea," *J. Amer. Heart Assoc.*, vol. 1, no. 4, Aug. 2012, Art. no. e000091.

[4] C. V. Senaratna, J. L. Perret, C. J. Lodge, A. J. Lowe, B. E. Campbell, M. C. Matheson, G. S. Hamilton, and S. C. Dharmage, "Prevalence of obstructive sleep apnea in the general population: A systematic review," *Sleep Med. Rev.*, vol. 34, pp. 70–81, Aug. 2017.

[5] L. Simpson, D. R. Hillman, M. N. Cooper, K. L. Ward, M. Hunter, S. Cullen, A. James, L. J. Palmer, S. Mukherjee, and P. Eastwood, "High prevalence of undiagnosed obstructive sleep apnoea in the general population and methods for screening for representative controls," *Sleep Breathing*, vol. 17, no. 3, pp. 967–973, Sep. 2013.

[6] R. B. Berry, R. Brooks, C. El Gamaldo, S. M. Harding, C. L. Marcus, and B. V. Vaughn, "The AASM manual for the scoring of sleep and associated events," in *Rules Terminology and Technical Specifications*, vol. 176. Darien, IL, USA: American Academy of Sleep Medicine, 2012.

[7] M. Ahmed, N. P. Patel, and I. Rosen, "Portable monitors in the diagnosis of obstructive sleep apnea," *Chest*, vol. 132, no. 5, pp. 1672–1677, Nov. 2007.

[8] A. T. Thornton, P. Singh, W. R. Ruehland, and P. D. Rochford, "AASM criteria for scoring respiratory events: Interaction between apnea sensor and hypopnea definition," *Sleep*, vol. 35, no. 3, pp. 425–432, Mar. 2012.

[9] F. Lofaso, A. Coste, M. P. D'Ortho, F. Zerah-Lancner, C. Delclaux, F. Goldenberg, and A. Harf, "Nasal obstruction as a risk factor for sleep apnoea syndrome," *Eur. Respiratory J.*, vol. 16, no. 4, pp. 639–643, Oct. 2000.

[10] I. Koutsourelakis, E. Vagiakis, C. Roussos, and S. Zakynthinos, "Obstructive sleep apnoea and oral breathing in patients free of nasal obstruction," *Eur. Respiratory J.*, vol. 28, no. 6, pp. 1222–1228, Dec. 2006.

[11] S. H. Lee, J. H. Choi, C. Shin, H. M. Lee, S. Y. Kwon, and S. H. Lee, "How does open-mouth breathing influence upper airway anatomy," *Laryngoscope*, vol. 117, no. 6, pp. 1102–1106, Jun. 2007.

[12] M. Friedman, H. Tanyeri, J. W. Lim, R. Landsberg, K. Vaidyanathan, and D. Caldarelli, "Effect of improved nasal breathing on obstructive sleep apnea," *Otolaryngology Head Neck Surg.*, vol. 122, no. 1, pp. 71–74, Jan. 2000.

[13] T. Sugiura. A. Noda, S. Nakata, Y. Yasuda, T. Soga, S. Miyata, and Y. Koike, "Influence of nasal resistance on initial acceptance of continuous positive airway pressure in treatment for obstructive sleep apnea syndrome," *Respiration*, vol. 74, no. 1, pp. 56–60, 2007.

[14] N. A. Dewan, F. J. Nieto, and V. K. Somers, "Intermittent hypoxemia and OSA: Implications for comorbidities," *Chest*, vol. 147, no. 1, pp. 266–274, Jan. 2015.

[15] D. Barak-Shinar, Y. Amos, and R. K. Bogan, "Sleep disordered breathing analysis in a general population using standard pulse oximeter signals," *Sleep Breathing*, vol. 17, no. 3, pp. 1109–1115, Sep. 2013.

[16] A. Garde, P. Dehkordi, W. Karlen, D. Wensley, J. M. Ansermino, and G. A. Dumont, "Development of a screening tool for sleep disordered breathing in children using the phone oximeter," *PLoS ONE*, vol. 9, no. 11, Nov. 2014, Art. no. e112959.

[17] C. Jonas, S. Thavagnanam, G. Blecher, G. Thambipillay, and A. Y. Teng, "Comparison of nocturnal pulse oximetry with polysomnography in children with sleep disordered breathing," *Sleep Breathing*, May 2019. doi: 10.1007/s11325-019-01861-z.

[18] D. Pevernagie, R. M. Aarts, and M. De Meyer, "The acoustics of snoring," *Sleep Med. Rev.*, vol. 14, no. 2, pp. 131–144, Apr. 2010.

[19] J. A. Fiz, R. Jané, J. Solà-Soler, J. Abad, M. Á. García, and J. Morera, "Continuous analysis and monitoring of snores and their relationship to the apnea-hypopnea index," *Laryngoscope*, vol. 120, no. 4, pp. 854–862, Apr. 2010.

[20] H. Jin, L. A. Lee, L. Song, Y. Li, J. Peng, N. Zhong, and X. Zhang, "Acoustic analysis of snoring in the diagnosis of obstructive sleep apnea syndrome: A call for more rigorous studies," *J. Clin. Sleep Med.*, vol. 11, no. 7, pp. 765–771, Jul. 2015.

[21] R. Jané, J. A. Fiz, J. Solà-Soler, J. Mesquita, and J. Morera, "Snoring analysis for the screening of sleep apnea hypopnea syndrome with a single-channel device developed using polysomnographic and snoring databases," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2011, pp. 8331–8333.
[22] E. Fino and M. Mazzetti, "Monitoring healthy and disturbed sleep through smartphone applications: A review of experimental evidence," *Sleep Breathing*, vol. 23, no. 1, pp. 13–24, Mar. 2019.
[23] A. Garde, P. Dehkordi, D. Wensley, J. M. Ansermino, and G. A. Dumont, "Pulse oximetry recorded from the Phone Oximeter for detection of obstructive sleep apnea events with and without oxygen desaturation in children," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 7692–7695.
[24] H. Nakano, K. Hirayama, Y. Sadamitsu, A. Toshimitsu, H. Fujita, S. Shin, and T. Tanigawa, "Monitoring sound to quantify snoring and sleep apnea severity using a smartphone: Proof of concept," *J. Clin. Sleep Med.*, vol. 10, no. 1, pp. 73–78, Jan. 2014.
[25] M. A. Cámara, Y. Castillo, D. Blanco-Almazán, L. Estrada, and R. Jané, "mHealth tools for monitoring Obstructive Sleep Apnea patients at home: Proof-of-concept," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 1555–1558.
[26] Y. Castillo, M. A. Cámara, D. Blanco-Almázan, and R. Jané, "Characterization of microphones for snoring and breathing events analysis in mHealth," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 1547–1550.
[27] M. Al-Mardini, F. Aloul, A. Sagahyroon, and L. Al-Husseini, "Classifying obstructive sleep apnea using smartphones," *J. Biomed. Inform.*, vol. 52, pp. 251–259, Dec. 2014.
[28] J. Behar, A. Roebuck, M. Shahid, J. Daly, A. Hallack, N. Palmius, J. Stradling, and G. D. Clifford, "SleepAp: An automated obstructive sleep apnoea screening application for smartphones," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 1, pp. 325–331, Jan. 2015.
[29] M. K. Guul, P. Jennum, and H. B. D. Sorensen, "Portable prescreening system for sleep apnea," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 4917–4920.
[30] I. Ferrer-Lluis, Y. Castillo-Escario, J. M. Montserrat, and R. Jané, "Automatic event detector from smartphone accelerometry: Pilot mHealth study for obstructive sleep apnea monitoring at home," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 4990–4993.
[31] Y. Castillo-Escario, I. Ferrer-Lluis, J. M. Montserrat, and R. Jané, "Automatic silence events detector from smartphone audio signals: A pilot mHealth system for sleep apnea monitoring at home," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 4982–4985.
[32] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Amer. J. Physiol. Heart Circulatory Physiol.*, vol. 278, no. 6, pp. H2039–H2049, Jun. 2000.
[33] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman, "Sample entropy analysis of neonatal heart rate variability," *Amer. J. Physiol. Heart Circulatory Physiol*, vol. 283, no. 3, pp. R789–R797, Sep. 2002.
[34] L. I. Lin, "A concordance correlation coefficient to evaluate reproducibility," *Biometrics*, vol. 45, no. 1, pp. 255–268, Mar. 1989.
[35] C. Plapous, C. Marro, and P. Scalart, "Improved signal-to-noise ratio estimation for speech enhancement," *IEEE Trans. Audio, Speech, Language Process.*, vol. 14, no. 6, pp. 2098–2108, Nov. 2006.
[36] L. Estrada, A. Torres, L. Sarlabous, and R. Jané, "Improvement in neural respiratory drive estimation from diaphragm electromyographic signals using fixed sample entropy," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 2, pp. 476–485, Mar. 2016.
[37] A. Roebuck and G. D. Clifford, "Comparison of standard and novel signal analysis approaches to obstructive sleep apnea classification," *Frontiers Bioeng. Biotechnol.*, vol. 3, p. 114, Aug. 2015.
[38] N. Esmaeili,H. Rabbani, S. Makaremi, M. Golabbakhsh, M. Saghaei, M. Parviz, and K. Naghibi, "Tracheal sound analysis for automatic detection of respiratory depression in adult patients during cataract surgery under sedation," *J. Med. Signals Sens.*, vol. 8, no. 3, pp. 140–146, Jul./Sep. 2018.
[39] A. J. Block, P. G. Boysen, J. W. Wynne, and L. A. Hunt, "Sleep apnea, hypopnea and oxygen desaturation in normal subjects—A strong male predominance," *New England J. Med.*, vol. 300, no. 10, pp. 513–517, Mar. 1979.
[40] F. Sériès, Y. Cormier, and J. La Forge, "Influence of apnea type and sleep stage on nocturnal postapneic desaturation," *Amer. Rev. Respiratory Disease*, vol. 141, no. 6, pp. 1522–1526, Jun. 1990.
[41] A. Kulkas, B. Duce, T. Leppänen, C. Hukins, and J. Töyräs, "Severity of desaturation events differs between hypopnea and obstructive apnea events and is modulated by their duration in obstructive sleep apnea," *Sleep Breathing*, vol. 21, no. 4, pp. 829–835, Dec. 2017.
[42] Q. A. Shamim-Uzzaman, S. Singh, and S. Chowdhuri, "Hypopnea definitions, determinants and dilemmas: A focused review," *Sleep Sci. Pract.*, vol. 2, no. 1, p. 7, Dec. 2018.
[43] Y. Jefferson, "Mouth breathing: Adverse effects on facial growth, health, academics, and behavior," *Gen. Dent.*, vol. 58, no. 1, pp. 18–25, Jan./Feb. 2010.

YOLANDA CASTILLO-ESCARIO received the B.S. degree in biomedical engineering from the Universitat de Barcelona (UB), Barcelona, Spain, in 2015, and the M.S. degree in biomedical engineering, specializing in bioinformatics and biomedical image and signal processing, from UB and the Universitat Politècnica de Catalunya (UPC), Barcelona, in 2016, where she is currently pursuing the Ph.D. degree in biomedical engineering with the Institute for Bioengineering of Catalonia (IBEC), Barcelona Institute of Science and Technology (BIST), Barcelona.

Since 2015, she has been a member of the Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Biomedicina (CIBER-BBN). From 2016 to 2018, she was a Research Assistant with the Biomedical Signal Processing and Interpretation Group, IBEC. In 2017, she was a Visiting International Research Student with Electrical and Computer Engineering in Medicine Group, University of British Columbia (UBC), and also with the Pediatric Anesthesia Research Team, BC Children's Hospital, Vancouver, Canada. Her research interests include biomedical signal processing and interpretation, mHealth systems, sleep apnea monitoring, motor control, and neuromuscular disorders.

IGNASI FERRER-LLUIS received the B.S. degree in biomedical engineering from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 2014, and the M.S. degree in biomedical engineering, specializing in robotics, bioinformatics, and biomedical image and signal processing from the Universitat de Barcelona (UB) and UPC, Barcelona, in 2015, where he is currently pursuing the Ph.D. degree in biomedical engineering with the Institute for Bioengineering of Catalonia (IBEC), Barcelona Institute of Science and Technology (BIST), Barcelona.

From 2015 to 2017, he was an early stage Researcher with the European Project List_MAPS, developing bioinformatic tools to visualize genomic data at GenXPRO GmbH, Frankfurt, Germany. In 2017, he was a Visiting Researcher with the Mathématiques et Informatique Appliquées du Génome à l'Environnement (MaIAGE) Group, Institut National de la Recherche Agronomique (INRA), Jouy en Josas, France. Since 2018, he has been a member of the Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Biomedicina (CIBER-BBN). His research interests include biomedical signal processing and interpretation, data science, mHealth technology, and obstructive sleep apnea monitoring and management.

JOSEP MARIA MONTSERRAT received the Physician degree, and the Ph.D. degree from the Universitat de Barcelona (UB), Barcelona, Spain, in 1974 and 1983, respectively. He is currently the Director of the Sleep Laboratory, Hospital Clinic, Barcelona, and a Principal Investigator of the Centro de Investigación Biomédica en Red de Enfermedades Respiratorias.

In 1979, he was a Visiting Researcher with New Cross Hospital, London, U.K., and in 1992, he was a Visiting Professor with McGill University, Montreal, Canada. Since 1995, he has been a Senior Consultant with Hospital Clínic, Barcelona, and since 2007, he has been a Professor with UB. He has authored more than 200 articles, and an H-index of 43. His research interests include technological and telemedicine studies (particularly CPAP and Bench studies), multicentric sleep studies, and basic sleep studies with animal models (particularly hypoxia/normoxia).

Prof. Montserrat is the Vice President of the Spanish Sleep Society (SES) and the President of the Spanish Respiratory Society.

RAIMON JANÉ (M'91–SM'14) received the Ph.D. degree from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 1989.

Since 2008, he has been the Principal Investigator of the Biomedical Signals and Systems (SISBIO) Group and a member of the Steering Committee of the Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBERBBN). He is currently the Director of Research with the Department of Automatic Control (ESAII), UPC, and the Scientific Group Leader of the Biomedical Signal Processing and Interpretation Group, Institute for Bioengineering of Catalonia (IBEC), Barcelona Institute of Science and Technology (BIST), Barcelona. He is also a Professor in the master's degree program and is the Coordinator of the Ph.D. program in biomedical engineering. His research interests include multimodal and multiscale biomedical signal processing in cardiorespiratory diseases and sleep disorders.

He is currently a member of the IEEE EMBS Technical Committee on Cardiopulmonary Systems. In 2005, he received the Barcelona City Prize from the Barcelona City Council for technology research. He is an Associate Editor of the Cardiovascular and Respiratory Systems Engineering Theme of the IEEE EMBC. He has been the President of the Spanish Society of Biomedical Engineering (SEIB), since 2012.