IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received April 23, 2021, accepted May 2, 2021, date of publication May 14, 2021, date of current version May 21, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3080258

Three-Stage Breathing Effort Quantification for Obstructive Sleep Apnea Detection Based on Thoracic and Abdominal Movement Signals

MUHAMMAD SHAUFIL ADHA[1] **AND TOMOHIKO IGASAKI**[2]
[1]Department of Advanced Industrial Science, Graduate School of Science and Technology, Kumamoto University, Kumamoto 860-8555, Japan
[2]Division of Informatics and Energy, Faculty of Advanced Science and Technology, Kumamoto University, Kumamoto 860-8555, Japan

Corresponding author: Tomohiko Igasaki (iga@cs.kumamoto-u.ac.jp)

This work involved human subjects or animals in its research. The authors confirm that all human/animal subject research procedures and protocols are exempt from review board approval.

ABSTRACT A level-IV home-based sleep apnea monitoring system that utilizes alternative sensors, such as dual respiratory inductance plethysmography (RIP) belts, is proposed to promote routine apnea monitoring. Notably, continued excursion may occur in RIP belt signals, owing to the imperfect relationship between thoracic and abdominal movements during obstructive events. Therefore, we propose a novel algorithm to detect obstructive apnea based on an obstructive reciprocal divergence (ORD) continued excursion model and to explore the possibility of a multistage breathing-effort evaluation model using only RIP signals. Using the developed approach, we detected obstructive sleep apnea with a high accuracy of $99.83 \pm 0.71\%$ and a slight reduction of $73.34 \pm 28.35\%$ in hypopnea performance with overall combined objective metrics of $89.38 \pm 10.53\%$. We found that introducing many stages improves specificity ($p < 0.001$). Furthermore, apart from apneic detection, we detected subtle changes in RIP signals qualitatively, which can help represent the inspiratory flow limitation (IFL) of the RIP. This study was validated by predicting an apnea hypopnea index (AHI) based on paradoxical breathing during sleep. A strong exponential relationship was observed between the proposed parameter based on the number of transitions with AHI ($R^2 = 0.98$; $p < 0.001$). The proposed approach can assist sleep technologists in characterizing obstructive and nonobstructive apneic events. Moreover, ORD is competent for further quantitative and qualitative IFL analysis and will significantly benefit the automated IFL detection system studies.

INDEX TERMS Abdominal signal, breathing effort, hypopnea, inspiratory flow limitation, obstructive sleep apnea, sleep-disordered breathing, thoracic signal.

I. INTRODUCTION

For a large majority of the population, adequate sleep helps inhibit a wide range of health complications. However, this is untrue for individuals that experience sleep-disordered breathing (SDB). The most common type of SDB is the intermittent cessation of breathing known as sleep apnea syndrome (SAS), and it includes obstructive sleep apnea (OSA) and the less severe case of hypopnea. Although SAS is not considered a life-threatening condition, considerable evidence linking SAS to cardiovascular and cerebrovascular diseases that can ultimately lead to a silent death has been reported [1]–[4]. The increased prevalence of SAS today—mainly promoted by the unhealthy dietary regimen—is alarming; however, it continues to remain one of the main under-diagnosed sleep disorders.

The most common and effective approach for SAS diagnosis is multichannel recording during night sleep or a polysomnography (PSG) assessment. Despite being considered the gold standard, public opinion regarding PSG is inconsistent. Patients are either financially reluctant and/or unwilling to spend an overnight session in a dedicated sleep laboratory [5]. This further intensifies the concurrent shortage of sleep laboratory equipment and sleep technologists (ST). Therefore, a level-IV (continuous single- or dual-biosensor recording) home-based sleep apnea monitoring system has been proposed for expediting diagnosis, and it has gained tremendous support [6], [7]. In compliance with

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

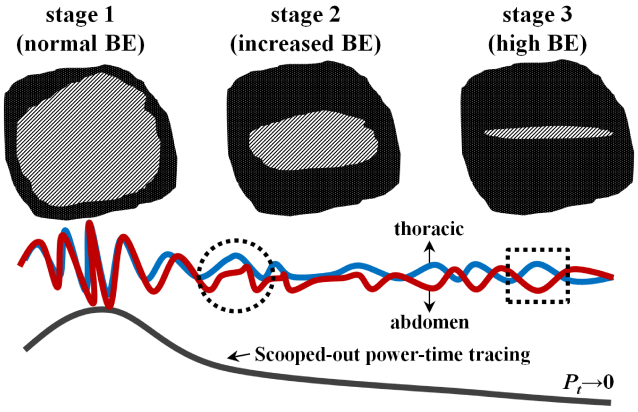

FIGURE 1. Conceptual diagram of the progressive behavior of the UA in OSA. The black and grey regions represent concealed pharyngeal structures (such as soft palate and epiglottis) and cross-sectional views of the trachea straight to the lung, respectively. The dashed circle indicates a chair-liked characteristic, and the dashed box highlights the phase difference between thoracic (blue line) and abdomen (red line) signals. Each is a renowned noninvasive pathophysiological hallmark of flow limitation and obstruction of the upper airway during sleep, respectively. The black line indicates a continued excursion of the absolute power of the sum of thoracic and abdominal signals (refer to Eq. (2)).

this, the American Academy of Sleep Medicine (AASM) task force has classified such respiratory sensors as *recommended* or *alternative* [8]. Recommended sensors have garnered serious consideration and are considered adequate for critical case findings; alternative sensors are yet to receive proper recognition within the automated detection or monitoring system framework. The explanation is simple: to detect a breathing-related disease, the sensors must be placed close to the anterior nares (or nostril) and the mouth to allow direct measurement of nasal and oral airflow, respectively. This is not the case for indirect sensors, such as the dual respiratory inductance plethysmography (RIP) belt that comprises thoracic and abdominal sections. This equipment makes routine respiratory event monitoring easy. Even with PSG assessment, RIP is considered for SAS scoring if the signal quality from the recommended sensors is unreliable or in the event of failure [8].

The rationale behind developing a dual-RIP-based automatic apnea detection (RAAD) system is justified by the clear pathophysiological connection between thoracic and abdominal movements, as illustrated in Fig. 1. Previous studies have frequently shown that the minimal amplitude excursion of the sum of thoracic and abdominal signals (RIPsum) is commonly observed during apnea [9]–[11]. However, a continued excursion in the power transformation of the RIPsum during OSA can occur when the summation does not produce a zero value [8], which consequently displays a concave characteristic, as illustrated in Fig. 1. The use of calibrated RIPsum signals can minimize this effect; however, in one study measuring tidal volume during sleep, even the calibrated RIPsum did not remain stable, owing to belt distortion or changes in body position [12]. It is technically difficult and impractical to perform RIP calibration with routine clinical PSGs [13], [14], which has led to previous RAAD studies disregarding the calibration procedure entirely [15]–[18]. Furthermore, several studies have suggested that an un-calibrated RIPsum is effective for the noninvasive evaluation of upper airway (UA) resistance (i.e., breathing effort (BE)), owing to the constrained airflow into the lung [19]. Moreover, it was suggested that the deflection in the un-calibrated RIPsum may provide a surrogate estimation of a relative change in the tidal volume when properly compared to the baseline condition [9], [20], [21]. Other studies have suggested that an airflow limitation can be inferred from RIPsum signals [9], [19], [22], [23]. Despite being classified as an alternative counterpart, the RIPsum can provide an abundance of clinical information in addition to apnea detection. There is sufficient evidence to suggest that the RIPsum can be applied to everyday use to help improve clinical diagnosis decisions.

The main propositions of this study are as follows:

- Although the continued excursion of the un-calibrated RIPsum may lead to unfavorable outcomes, the observed patterns can still explain breathing dynamics if modeled properly. Herein, we demonstrate that it is possible to develop a mathematical model of continued excursion using the un-calibrated RIPsum signals for RAAD application.
- If the above proposition is feasible, introducing an intermediate stage between normal breathing and apnea events is possible. The signal excursion must drop below a certain percentage from a pre-baseline to indicate apnea. However, the continued excursion of RIP signals caused by an apnea event by no means drops below the threshold immediately after the pre-baseline is defined. The signals must transit at a certain level before they further drop or remain at the current level. We hypothesize that this intermediate stage may provide additional information on the RIP signals. Further, a pre-baseline (also known as pre-event baseline) is referred to a baseline condition prior to an apnea event. That is, each event is associated with a unique baseline value that further complicates automatic detection.

These issues have never been considered in published RAAD algorithms. The novelty of our study is as follows:

- To the best of our knowledge, this is the first study that attempts to introduce an intermediate stage between normal breathing and apnea-related breathing using RIPsum signals. This study develops a multistage BE for automatic apnea event detection.
- To this end, this study introduces new obstructive reciprocal divergence (ORD) mechanics to model the continued excursion in RIPsum signals.

Our main objective is to develop an algorithm for the RAAD system that can detect apnea events and concomitantly deliver important clinical information related to breathing for improved diagnosis. In this study, we consider the analysis performed on only RIP thoracic and abdominal signals (i.e., dual-RAAD system). The proposed method is

expected to improve the acquisition of noninvasive information with a minimum number of PSG channels. This will not only improve our understanding of the dynamic behavior of the UA, but it will also promote more practical real-time solutions.

The remainder of this paper is organized as follows. Section II details the literature regarding the differences among algorithms used in current RAAD studies. Section III details the mathematical modeling of the continued excursion and proposes a multistage classification approach. It also describes the validation procedure of the proposed algorithm. Section IV reports the results and presents a comprehensive discussion. Finally, Section V provides concluding remarks and discusses the limitations of this work with future research recommendations.

II. RELATED WORK

Although automatic SAS detection algorithms have received considerable attention [24], the development of different algorithms for RAAD systems is rare. Because the level-IV system includes single- or dual- bio-sensor recordings, algorithms use either thoracic, abdominal, or both belt signals (RIPsum). For example, one study developed an OSA detection algorithm that adopts an adaptive envelope-tracking function following the amplitude excursion of a thoracic signal [16]. The tracking function is adaptive in that for each breath, a new threshold is defined based on the height of the previous peak. The algorithm was designed to exploit an amplitude excursion rate, which, for an OSA event, occurs at a slower rate than normal respiration. Therefore, it induces apparent changes in the envelope. The results are significantly correlated with manual scoring with R-square values of 0.73 and 0.55 for training and validation sets, respectively. However, hypopnea events were not considered in the proposed work. In another study, an algorithm based on the adaptive threshold of the power transformation of the RIPsum was applied to detect both OSA and hypopnea events [17]. That is, a threshold was defined for every power segment calculated dynamically from 120 s preceding the current segment. Therefore, any observed power drop below a current threshold is considered an OSA or hypopnea event with no distinction between the two.

Lin *et al.* applied support vector machine to classify three categories: normal, OSA, and central sleep apnea (CSA) using novel parameters extracted from an adaptive nonharmonic model that simulates the thoracic and abdominal signals [15]. Although the detection of CSA was considered, further detection was not conducted for hypopnea events. An algorithm developed by Steenkiste *et al.* attempted to detect apnea based on a deep-learning neural network using raw thoracic and abdominal signals. They claimed that the use of deep learning for blind feature extraction is more suitable than human-engineered features that may miss important apnea information [18]. However, there is no distinction between OSA and hypopnea events. Moreover, it is evident that the infamous undisclosed nature of deep learning nearly prohibits further qualitative analysis, which could be equally important. Even under these conditions, important information associated with RIP signals that change during an apnea event are not shown.

Current RAAD algorithms developed on machine-learning platforms suffer from several limitations. First, a classifier based on machine learning can be considered a highly complex black-box system that can be considered successful if it maintains high detection performance. However, the underlying mechanics is often inaccessible, owing to its calculation complexity. Thus, in the event of an apnea, information such as a percentage drop from the pre-baseline is not known, because it is not considered in the learning process. It may be inadvertently considered, but the information is not made transparent by the machine. This leads to the incapability of distinguishing apnea and hypopnea events. Identifying this information is crucial, because it supports apnea annotation guidelines. Second, it is well-established that machine learning, especially a neural-network approach as in [18], is highly computational expensive and can impose additional drawbacks for real-time application.

Current algorithms applied for RAAD systems cannot distinguish between OSA and hypopnea events. No algorithm has attempted to perform apnea detection concomitantly and analyze the qualitative changes in RIP during an apnea event. Few mainstream studies have emphasized this necessity, especially those that implement alternative sensors.

III. METHOD

A. DATABASE

The data of 25 adult patients with a possible diagnosis of SDB were randomly selected from the sleep apnea database publicly available at St. Vincent's University Hospital/University College Dublin [25]. The patients' demographics are summarized in Table 1. The patients did not have known autonomic dysfunctions (e.g., multiple sclerosis) or cardiac disease (e.g., arrhythmias). In addition, the patients did not take medications that would interfere with the heart rate prior to the PSG assessment. The annotation of apnea onset and duration was conducted by an experienced ST.

B. PRE-PROCESSING

Given the wide range of sensor characteristics used in PSG studies and available commercially, linearization is highly encouraged for reproducibility. Therefore,

$$Y = \frac{4y}{y_{max}}, \tag{1}$$

where vector Y is derived from the linearization of any physiological signal of y, and $y_{\max}$ denotes the cut-off amplitude of the sensor. The thoracic and abdominal signals from the RIP belts were bandpass-filtered between the frequencies, 0.07–0.8 Hz, including the extreme breathing rate, which can be between 4–48 beats per minute [17], [26]. Each signal was segmented into a 5-s interval with a 90% overlap. The current segment contains 4.5 s of data from the previous

TABLE 1. Subject Demographics, Event Classification, and AHI Characterization.

Sex	21 Male, 4 Female (N = 25)
Age	50 ± 10 years; Range: 28–68 years
Body mass index (BMI)	31.6 ± 4.0 kg/m²
Apnea hypopnea index (AHI)	24.1 ± 20.3; Range 1.7–90.9
Apneas Hypopneas	Obstructive, Central, and Mixed
AHI characterization	<5 Normal; 6–15 Mild; 16–30 Moderate; ≥30 Severe

Data are shown as mean ± standard deviation. AHI = "apnea hypopnea index."

segment. Considering such an overlapped segment is a strategy commonly employed in physiological signal processing for real-time applications. This approach allows the algorithm to perform calculations on a per-second basis (preferably faster), which increases the granularity of the detection [18]. A revised-absolute power expression of RIPsum was applied to quantify the BE; it is expressed as

$$e_{i,j} = \left|tho_{i,j} + abd_{i,j}\right|^2,$$
$$P_j = \frac{1}{N}\sum_{i=1}^{N} e_{i,j}, \quad 0 \le e_{i,j} \le \left(\mu_{e_{i,j}} + \sigma_{e_{i,j}}\right). \tag{2}$$

Here, $e_{i,j}$ and P represent the i^{th} elements of the j^{th} segment and the power of the RIPsum per segment, respectively. One immediate benefit of applying the revised expression is preventing outliers caused by distorted belt tension. As reported in a previous study [15], absolute power is unsuitable for OSA detection; therefore, a ratio of the upper quartile amplitudes between two consecutive segments (current and previous) for both thoracic and abdominal signals was proposed for outlier exclusion. Herein, with a minor revision of the absolute power calculation, it is demonstrated that OSA can be detected with acceptable accuracy.

C. PARAMETER DESIGN

In all cases, the patency of the UA is highly unstable, and it has a higher tendency to collapse as patients sleep. The degree of collapse presents several variations: it might be fully opened, narrowed to some extent, or completely blocked. Further, it may also behave erratically under these conditions. A previous study demonstrated that obstructions may develop from the collapse of either one or any combination of pharyngeal structures [27], [28]. However, for simplification, a cross-section of the UA was considered, as illustrated in Fig. 1. As the patient struggles to breathe through the fully (or nearly) collapsed UA at stage 3, a thoracoabdominal paradoxical movement becomes apparent. In the ideal case of no synchronization across the event, the remaining excursion cancels each other. Although this is theoretically true, realistically, as P approaches zero (black line in Fig. 1), a characteristic scooped-out (or concave) trend is observed.

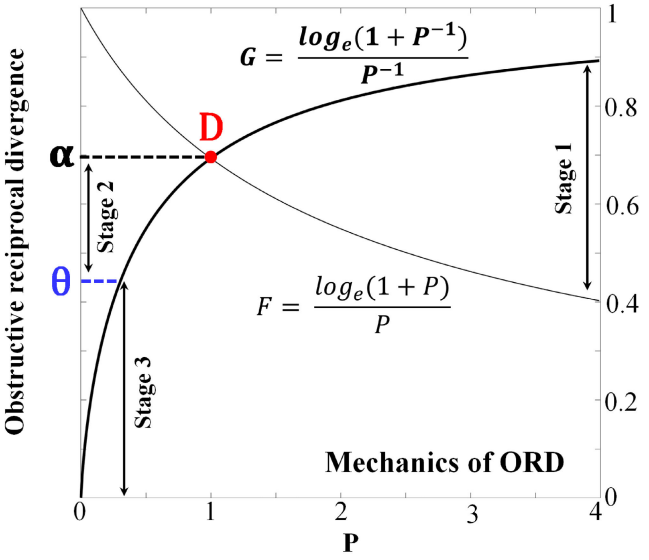

FIGURE 2. Working mechanics of ORD, where *D* represents the divergence point between two functions. The abscissa represents the linearized input (refer to Eq. (1)), and the ordinate is ORD. Alpha (α) is a fixed threshold denoted by the black dotted line. Theta (θ) is an adaptive threshold denoted by the blue dotted line. The range of stages 2 and 3 change accordingly to the dynamic of breathing effort.

Thus,

$$\lim_{P\to 0} F_j = \frac{\log_e\left(P_j + 1\right)}{P_j}, \quad 0 < F_j < 1. \tag{3}$$

Although a sudden collapse is possible, the expression still applies, and it provides further insight into relevant physiological information. As dictated by the AASM task force, OSA is measured from the nadir preceding the first breath reduced from the baseline. The concept of ORD is introduced given the difficulty in measuring the baseline and finding the approximate location of a nadir. That is, as P approaches zero, the multiplicative inverse of P diverges to infinity. The ORD of F is defined as

$$\lim_{P\to 0} G_j = \frac{\log_e\left((P_j)^{-1} + 1\right)}{(P_j)^{-1}}, \quad 0 < G_j < 1. \tag{4}$$

The terms, F and G, do not present any abbreviation; however, the range of both parameters is bounded within zero to one. Hence, they are defined as the normalized powers of P and a reciprocal of P, respectively. Equations (3) and (4) have unique solutions; therefore, F and G should always intercept at the alpha (α) value (refer to Fig. 2). However, for simplicity,

$$D_k = F_{j-1} < G_{j-1} \text{ and } F_j > G_j. \tag{5}$$

The divergence point (D) suggests that the BE starts to increase considerably, and index k represents the number of occurrences that satisfy Eq. (5). The theoretical working mechanics of ORD are presented in Fig. 2. Here, P approaches zero to the left; however, in a real application, P approaches zero following the time course of the RIPsum (to the right). This work does not consider both events annotated with apnea and hypopnea central (Table 1).

D. CLASSIFICATION

1) BREATHING EFFORT CLASSIFICATION

The un-calibrated RIPsum is effective for the noninvasive evaluation of the BE [19], [29]. Because the ORD relies heavily on the paradoxical movement between the thoracic and abdominal belts, the use of the BE term is deemed more suitable. A key difference between the non-apneic-associated BE and apneic-associated BE is that the latter displays considerably more apparent changes in oxygen desaturation several seconds after breathing has ceased. To distinguish between the two, BE is defined as

$$BE_j = \begin{cases} \text{stage 1 (S1)}, & F_j \leq G_j \\ \text{stage 2 (S2)}, & F_j > G_j \text{ and } \theta_k \leq G_j \leq \alpha \\ \text{stage 3 (S3)}, & F_j > G_j \text{ and } G_j < \theta_k, \end{cases} \quad (6)$$

where theta (θ) is a D-adaptive threshold defined as

$$\theta_k = \left(\mathfrak{m}_{R(G120)} \times \sim \frac{T}{100} \right) + \mathfrak{m}_{R(G120)}. \quad (7)$$

Here, T and $G120$ represent a fixed threshold in a unit of percentage, such as T $<$ 0, and 120 s preceding D of parameter G, respectively. Further, $\mathrm{T}_{R(\mathrm{G120})}$ represents a sample median (a revised $G120$), where R denotes the revised function. It is not necessary to apply the revision on $G120$ for the baseline calculation; however, the AASM Chicago Consensus highly encourages the consideration of an individual without a stable breathing pattern where the three largest breaths are selected within 120 s preceding an onset [15]. Thus, the revised function is expressed as

$$R(x) = \max_{Pc\%} [S(x)], \quad (8)$$

where S and x denote a sorting function and an input vector, respectively. A maximum of 20% (i.e., $Pc = 20$) was selected from the sorted input to resolve cases where there were several events within a baseline period [17]. However, our view on this follows the AASM Chicago definition. That is, 20% of the sorted input should be regarded as approximately the three largest breaths. Nevertheless, it is demonstrated that 20% is sufficient to perform a robust evaluation for both OSA and hypopnea. Let us consider an example to betterunderstand Eqs. (7) and (8). Assume that the first D (i.e., $k = 1$) is defined 300 s after a patient falls asleep. Therefore, $G120$ is simply a data vector of G within the period of 179–300 s. Then, we apply (8) in which the input vector, x, in this case is $G120$. For the example, let us define a pseudo vector of $G120$ that contains 10 data samples of G in the form $G120 = \langle 0.4, 0.5, 0.2, 0.1, 0.8, 0.7, 0.75, 0.85, 0.95, 0.3 \rangle$. Applying the sorting function, the sorted $G120$ takes the form $S(G120) = \langle 0.1, 0.2, 0.3, 0.4, 0.5, 0.7, 0.75, 0.8, 0.85, 0.95 \rangle$. Now, we consider only the maximum 20% of the sorted elements, i.e., $R(G120) = \langle 0.85, 0.95 \rangle$.

2) ACCESSING THRESHOLD USING PRIOR DISTRIBUTION

The AASM task force offers a comprehensive guideline on how to annotate a respiratory event as apnea or hypopnea. A respiratory event is annotated if i) there is a drop in the peak excursion by 90% (apnea) and 30% (hypopnea) of the pre-baseline; ii) the duration of percentage drop is $\geq$10 s; and iii) there is $\geq$4% oxygen desaturation from the pre-baseline. This fundamental information is applied by an experienced ST to annotate the onset time and duration of apneic events. In addition to the AASM guidelines, previous evidence suggests that respiratory events based on a desaturation of at least 4% are associated with increased risk of cardiovascular consequences [30]. Thus, we consider 4% in the analysis as the separation point to distinguish between apnea and non-apnea events. It is probably relevant to follow the definition of the task force, where $T = 90\%$ and 30% (consensus) or 50% (alternative) for OSA and hypopnea detection, respectively. However, because G is ORD-imposed, its characteristic changes compare slightly to that of P. Consequently, T is referred to as T_{ORD}. However, only a minor shift is expected. To determine T_{ORD}, each reduction of G from the baseline is quantified as

$$g_d\,(\%) = \frac{100 \times \sim \left(\mathfrak{m}_{G10} - \mathfrak{m}_{R(G120)} \right)}{\mathfrak{m}_{R(G120)}}, \quad (9)$$

where g_d denotes the percentage drop of G from the baseline, and index d is either $\geq$4% (OSA only) or $<$4%, representing a drop in oxygen desaturation. $\mathfrak{m}_{G10}$ represents the sample median of 10 s following D of G. To be categorized as OSA or hypopnea, the duration of the signal excursion drop must be $\geq$10 s. However, it is demonstrated that 10 s are adequate to analyze the fundamental behavior of apneic events. Furthermore, only OSA events are quantified for two principal bases. First, F and G are exclusively designed for extreme apnea cases (i.e., OSA), for which complete obstructions are consistently observed. Second, this serves as a validation method for hypopnea events, as an oxygen desaturation of $\geq$4% can occur even in a narrowed UA.

One of the many challenges of SAS studies is unbalanced datasets. In the present database, the total number of annotated events for hypopnea (N = 1,567) is four times that for apnea (N = 352); the number of non-apneic events is far larger. This difference poses considerable difficulty for studies that apply machine learning [18]. The number of elements per class must be equivalent to avoid bias in the trained model. No complex machine-leaning procedure is performed in this study; however, we considered the probability density function (p.d.f.) between two classes (S2 and S3) to define T_{ORD}. The histogram was first developed for $g_{\geq 4\%}$ and $g_{<4\%}$ and was Gaussian-fitted. Therefore, T_{ORD} can be defined as a mid-intersection between two fitted distributions. Fig. 3 shows the distribution for all subjects, wherein skewed distributions are observed. However, a Gaussian function was applied for the following two reasons: (i) In this study, the pattern related information or p.d.f. value itself is not used as part of the T_{ORD} definition. For example, if we consider spurious and skewed empirical distributions developed directly from each histogram as shown in Fig. 3 (blue line color), the T_{ORD} value that results from an intersection between these lines

TABLE 2. Performance Evaluation Criteria Based on Breathing Effort (BE) Transition.

BE Transition	Metric	Metric event	Physiological definition
Group 1: S1→S3; S2→S3;	*tp*	Marked by an ST as OSA/HYPOPNEA at the end of the transition.	The listed transition suggests that the UA is progressively obstructed, enough to cause ≥4% oxygen desaturation (i.e., the occurrence of paradoxical breathing).
	fn	Marked by an ST as NORMAL at the end of the transition.	
Group 2: S1→S1; S1→S2; S2→S2; S3→S1; S2→S1; S3→S2	*tn*	Marked by an ST as NORMAL at the end of the transition.	The listed transition suggests normal, narrowed, or recovery breathing. (i.e., non-related to paradoxical breathing).
	fp	Marked by an ST as OSA/HYPOPNEA at the end of the transition.	

Symbol "→" represents a transition

(blue-filled circle) is approximately equivalent to that of an intersection fitted by a Gaussian function (black-filled circle). Therefore, to determine the intersection point by using a different distribution function yields less significant changes in the results. (ii) We seek to determine a unique intersection between the fitted lines. It is possible to encounter multiple intersection points in a spurious line. Therefore, a smooth fitted line for each distribution is required to avoid such conditions and guarantee a unique mid-intersection point. In addition, a Gaussian fitted distribution can be considered a direct approach. Most distribution functions cannot be fitted directly with negative values; therefore, certain transformation calculations must be imposed on g_d, which can further complicate the direct interpretation of T_{ORD}. Another solution involves the application of the proper value of the bin-width interval that can slightly reduce the spurious condition. A longer interval may result in a significant loss of information. In contrast, a short interval may further degrade the spurious condition. In this study, 5% was empirically chosen based on our preliminary analysis. We demonstrate that the proposed approach is adequate for high-performance evaluation.

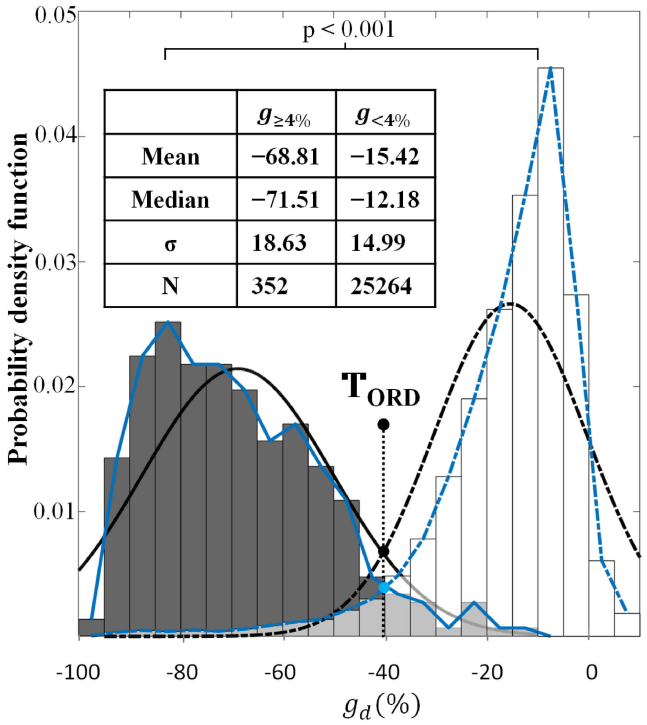

FIGURE 3. Black and white histograms for $g_{\geq 4\%}$ and $g_{<4\%}$, respectively. Gaussian-fitted curves for $g_{\geq 4\%}$ (solid black line) and $g_{<4\%}$ (dotted black line). The line developed directly from $g_{\geq 4\%}$ (solid blue line) and $g_{<4\%}$ (dotted blue line) is also shown. Statistical properties shown in the table are associated with the original histogram (not for fitting). For reproducibility, the bin widths of each histogram are regulated to an interval of 5%. Statistical comparison was conducted using Welch's t-test for unequal sample sizes.

3) PERFORMANCE EVALUATION BASED ON BREATHING TRANSITION

The proposed algorithm performs an event-based evaluation wherein metrics are associated with certain events [17]. In this study, we considered an adjacent transition of the BE stage as an event. For example, as illustrated in left panel of Fig. 4 (bottom plot), if the transition from S2 to S3 is encountered within the apnea period marked by an ST, then it will be counted as a true positive (*tp*). In contrast, it will be a false positive (*fp*) if the period is marked as *normal* (nonapnea). The right panel (bottom plot) illustrates the transition from S1 to S2 within the normal period; therefore, it is counted as a true negative (*tn*). In contrast, it will be counted as a false negative (*fn*) if it occurs within the apnea period marked by an ST. Please refer to Table 2 for the complete transition alongside its physiological definition.

The automated detection of apnea events may assist in diagnostic decision-making and may provide expedited screening; however, for real-time applications, most detection systems are developed to provide appropriate intervention before clinical symptoms appear. It is well-established that continuous cessation of breathing for 10 s results in a ≥4% drop in oxygen desaturation [8]. Therefore, we argue that the detection cue should be raised after the transition from either S1 or S2 to S3 is observed (without imposing a 10 s wait time to determine if it as an apnea, which can help prevent further drops in oxygen desaturation) to initiate a therapeutic device (i.e., early intervention). For example, one device that utilizes the early detection of abnormal physiological changes during sleep is the auto-adjusting positive airways pressure (APAP) device [31], [32]. Therefore, the transition S3→S3 is not evaluated assuming the intervention cue raised previously succeeds in preventing continuous UA collapse. In this study, S3 suggests that the UA has collapsed. Moreover, the event-based (i.e., transition-based) approach

TABLE 3. Performance for Three Different Thresholds (Averaged from 25 Iterations).

	Sensitivity (%)		Specificity (%)	Accuracy (%)	Combined objective (%)
	OSA	Hypopnea			
T = −30%	98.52 ± 4.07 n.s.	73.90 ± 27.38 n.s.	90.79 ± 4.41***	90.77 ± 4.31***	84.00 ± 10.02***
T = −50%	96.73 ± 5.20*	49.42 ± 25.33***	95.33 ± 2.52***	95.28 ± 2.45***	82.72 ± 9.98***
T_{ORD} = −40.75%	99.83 ± 0.71	73.34 ± 28.35	98.79 ± 0.92	98.77 ± 0.89	89.38 ± 10.53

Data are shown as mean ± standard deviation. Statistical comparison was conducted using Student's t-test between −30 and −50% with T_{ORD}.
n.s.: Not significant; *: p<0.05; ***: p<0.001.

provides more dynamic information on the breathing state than the temporal overlap. Performance indicators, such as the sensitivity (*SEN*), specificity (*SPE*), and accuracy (*ACC*), are respectively computed as

$$SEN = \frac{tp}{tp + fn}, \tag{10}$$

$$SPE = \frac{tn}{tn + fp}, \tag{11}$$

$$ACC = \frac{tp + tn}{tp + tn + fp + fn}. \tag{12}$$

The combined objective (*CO*) is encouraged for considering imbalance events [33] and is defined as

$$CO = \frac{\frac{SEN_{OSA} + SEN_{Hypoapnea}}{2} + SPE + ACC}{3}. \tag{13}$$

In addition, T_{ORD} is remeasured using the leave-one-subject-out cross-validation method to prevent over-fitting. All data from the 24 patients were utilized to measure T_{ORD} and evaluate the performance factors (i.e., *SEN*, *SPE*, *ACC*, and *CO*) of the patient that was not included. A total of 25 iterations were performed, and the evaluated performances for each patient were averaged. The leave-one-out cross-validation strategy offers an advantage wherein it can be considered as an inter-subject evaluation. Furthermore, it is considerably more robust compared with intra-subject evaluation (i.e., a threshold is measured with data collected from a subject and then the performance is evaluated using unobserved data from the same subject).

4) ORD FOR AHI PREDICTION

The common index used to characterize the severity of SAS is the AHI value, which is defined as the number of apnea and hypopnea events per hour of sleep. The prediction of the AHI value can help expedite screening before a complete manual diagnosis is conducted with PSG. This sub-section serves as a second-degree validation procedure that assumes the temporal information of annotated apneic events (i.e., onset and duration) except for the AHI value, which is unknown. Therefore, all observed paradoxical breathing, as seen by ORD (group-1 transitions), are considered. Nonrelated paradoxical breathing (group-2 transitions) re also considered (Table 2). Therefore,

$$\beta = \log_e\left(\frac{pb}{npb}\right), \tag{14}$$

where β is defined as the log transformation of the ratio between the total count of the observed paradoxical breathing (*pb*) and the total observed nonparadoxical breathing (*npb*). From a different viewpoint, high β values suggest increased paradoxical breathing during sleep, which can be associated with apnea occurrence in most cases. Low β values suggest a low occurrence of paradoxical breathing during night sleep. Paradoxical breathing can occur independently of obstructive apneic events.

IV. RESULTS AND DISCUSSION

A. ORD AND HYPOPNEA DETECTION

Table 3 presents the sensitivity for both OSA and hypopnea. A slight reduction in the hypopnea performance (73.34 ± 28.35%) compared with that of OSA (99.83 ± 0.71) was observed. This is consistent with the results of previous studies. For example, even with a more sophisticated algorithm, the reported results for hypopnea detection ranged around 70% although six physiological signals from PSG were applied [34]. Lin *et al.* reported that the inclusion of hypopnea downgraded the final analysis outcome, thus omitting the events entirely [15]. Even in a study related to respiration, hypopnea detection did not produce good agreement with manually annotated events largely because of the use of a thermistor airflow sensor in the PSG instead of a polyvinylidene fluoride film [35]. It is well-established that the excursion of the respiration signal amplitude for hypopnea events is not consistently visible [8]. With ORD, most unexplained hypopnea events remained at the second stage, owing to the low degree of paradoxical movement. Thus, the RIPsum does not precisely sum to zero (i.e., continued excursion does not occur). The typical behavior for failed hypopnea detection shown in Fig. 4 (right panel) suggests that approximately 70% of the events produce a clear continued excursion for a successfully detected hypopnea event.

B. THRESHOLD COMPARISON

Fig. 3 shows the distributions for $g_{\geq 4\%}$ and $g_{<4\%}$, which are extracted from all subjects. A significant difference is observed between each distribution ($p < 0.001$). The averaged T_{ORD} result from cross-validation is −40.75 ± 0.55%. Table 3 presents the cross-validated performance comparison between the three different thresholds. The overall evaluation of *CO* indicates that T_{ORD} (89.38 ± 10.53%) significantly outperforms that of 30% (84.00 ± 10.02%), followed by 50% (82.72 ± 09.98%). The significant drop considerably influences the lower *CO* for the 50% threshold in hypopnea detection (49.42 ± 25.33%;

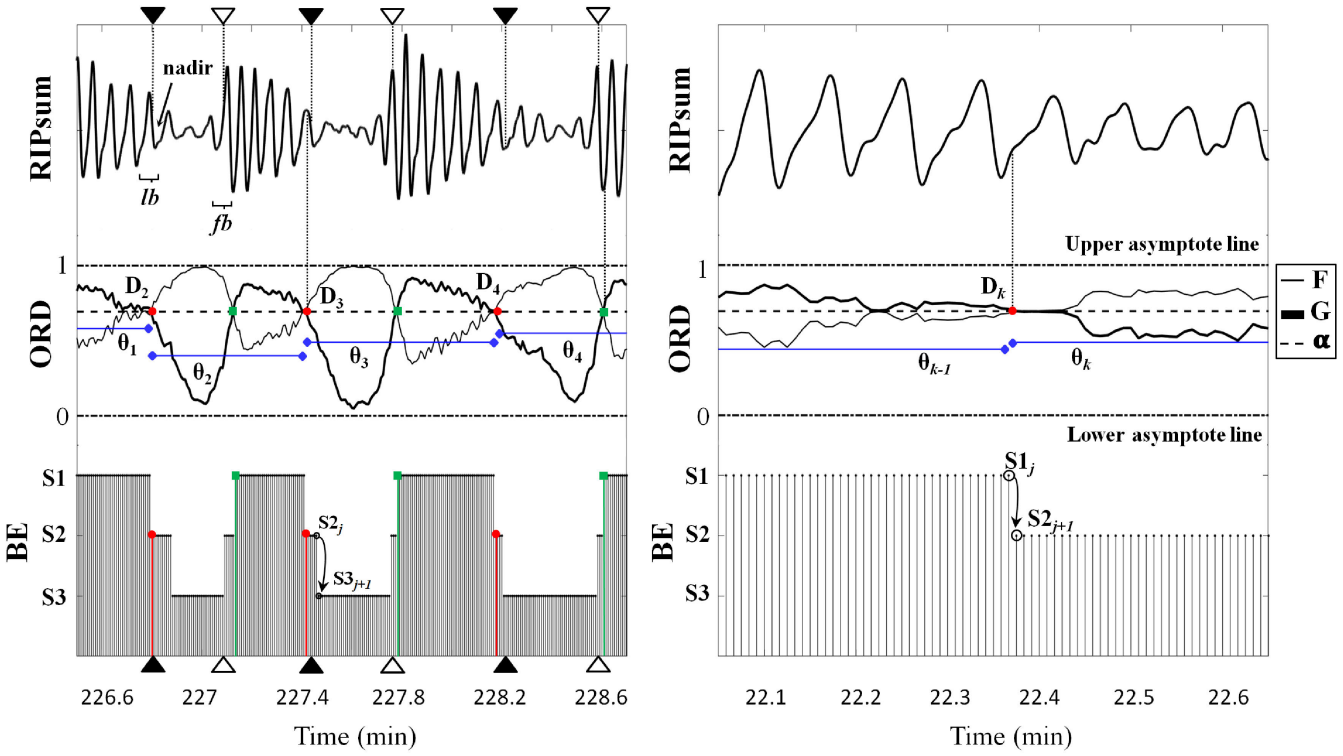

FIGURE 4. Sample tracing of RIPsum (top left plot), parameters F and G (middle left plot), and BE (bottom left plot) depicting OSA events. The black and white triangles represent OSA onset and offset, respectively; they are marked by an ST. The red-filled circle indicates *D*, which represents an approximate peak of the last breath (*lb*) in the event of an apnea. A green-filled square indicates the end of the first breath (*fb*). The diamond-edged blue line illustrates a theta for every respective *D*. A dotted line represents a constant value known as the alpha (α) threshold. The right panel illustrates sample tracing with the same order as the left panel, which indicates the transition from stage 1 to 2 and remains at the current stage. The progressive changes the inspiratory portion of the RIPsum signals. The ordinate of the left and right panel of the RIPsum is comparable.

$p < 0.001$). This result is consistent with the task-force recommendation: a 50% drop in flow is required for the single pediatric definition of hypopnea [8]. The task force redefined pediatric as all children < 18 years old, whereas the patients included in this study were above 18 years old (50 ± 10 years). Therefore, the 30% drop in the flow is recommended by consensus for adult patients. This explains the high *CO* for the 30% threshold compared with that for 50%. Moreover, the reason T_{ORD} is considerably more effective than the 30% threshold standard needs to be discussed. However, no significant difference ($p > 0.05$) was observed between T_{ORD} and the 30% threshold for both OSA and hypopnea events. The higher *CO* for T_{ORD} results from the elevated specificity (98.79 ± 0.92%) compared to that for the 30% threshold (90.79 ± 4.41%). This suggests T_{ORD} is considerably more robust for nonapnea event detection, although it maintains comparable sensitivity with the standard threshold for both OSA and hypopnea. However, the direct comparison of T_{ORD} with the 30% threshold (also for 50%) is not possible; therefore, redefining the standard threshold with T_{ORD} is highly irrelevant. This is because T_{ORD} is determined using *G*, which is transformed from the original RIPsum. Thus, we conclude that the slight shift of T_{ORD} from the standard threshold can be attributed to the changed properties of the parameter. The use of T_{ORD} with different databases is yet to be tested. Therefore, the generalization of the T_{ORD} is currently not possible.

C. PERFORMANCE COMPARISON BETWEEN RAAD-RELATED STUDIES

Table 4 summarizes the comparison results between several RAAD systems proposed previously. However, studies that applied the RIP combination with other physiological signals (e.g., SpO2 and respiration) were excluded. Considering the initial development of ORD, the use of a publicly available database with a relatively small number of patients (N = 25) is analytically appropriate. Furthermore, the same database was used by [17]. For all studies, separate sensitivities for OSA and hypopnea events were not reported. When comparing the standard performance metrics (*SEN*, *SPE*, and *ACC*), the ORD approach outperforms all previous methods. Although some studies proposed different functions to calculate standard metrics, the definition remains the same as the direct generalization of the binary categorical response data [15]. The study-dependent metric (*SDM*) was used when some standard metrics were not defined by the state-of-the-art approaches. Although the *SDM* between studies cannot be compared directly [33], the corresponding values are presented in Table 4 for the subjective evaluation. The SDM in this work corresponds to *CO*.

Based on the proposed segment overlapping setting for ORD, a single stage is defined for every 0.5 s (i.e., pre-segment = 120 s; segment = 5 s; and detection rate = 0.5 s/segment). The pre-segment was utilized in some studies to define a pre-baseline. It is also desirable to keep the

TABLE 4. Comparison of Several RAAD Studies using Either Thoracic or Abdominal Belt Signals or Both.

Studies	No. of subjects	Signals	Evaluation approach	Classifier	Pre-segment (s); Segment (s); Detection rate (s/segment)	SEN (%)			SPE (%)	ACC (%)	SDM (%)	AHI Prediction analysis (Yes/No)
						OSA	Hyp.	Combine				
[15][a]	34	RIPsum	80% training, 20% test, 25-fold cross validation	SVM	60; 10; 0.5	88.6 ± 9.06	—	—	79.4 ± 9.4	81.8 ± 9.4	84.0 ± 9.06	No
[16][b]	116	Thoracic	50% training, 50% test	Adaptive threshold	—; 30; 30	—	—	80	73.5	67	85	Yes
[17]	25	RIPsum	per subject test	Adaptive threshold	120; 6; 0.6	—	—	72	—	—	65.2	No
[18]	2100	Thoracic	5-fold cross-validation	LSTM-neural network	—; 30; 1	—	—	57.9 ± 8.6	73.9 ± 10	71.1 ± 6.8	71.5 ± 1.7	Yes
		Abdominal				—	—	62.9 ± 3.5	77.2 ± 4.5	74.7 ± 3.1	76.9 ± 0.8	

Data are shown as mean ± standard deviation; a: Calculation of accuracy and SDM including CSA events; b: For AHI > 15; SVM = Support vector machine; LSTM = Long short-term memory.

period short, as in [15]; however, this study adheres with the AASM Chicago guidelines, which recommends 120 s. It is also desirable to maintain a lower detection rate while ensuring high detection performance for a speedy and accurate intervention. This rate is comparable with that of [15], and it is 50% lower when compared with that of [18]. Thus, a high performance can be obtained even with a multistage approach and a lower detection rate. Thus, it can be considered to be an additional merit to the ORD algorithm. We do not need to sacrifice a longer segment to obtain such performances. Indeed, we did not consider the computational time for each segment when they commonly considered a real-time setting. Furthermore, in this study, no comparison analysis was conducted between the ORD algorithm and the machine-learning method. However, hypothetically, the real-time implementation of the ORD algorithm may have a lower overall computational time and use fewer resources than those developed on machine-learning platforms.

D. AHI PREDICTION BASED ON ORD TRANSITION

We predicted the AHI based only on the number of paradoxical breathing incidents to further validate the proposed approach. Fig. 5 presents the fitted line of β against AHI. The AHI value in the left panel (with central events) is higher compared with the values in the right panel (without central events). A reduced relationship strength is observed ($R^2 = 0.70$) compared with the fitted line on the right panel. The transitions for central events were not considered to predict the associated errors. Additionally, a strong exponential relationship ($R^2 = 0.98$) of β against the adjusted AHI was observed. The result suggests that it is possible to predict the AHI given only the number of observed breathing transitions without knowing the temporal apneic information.

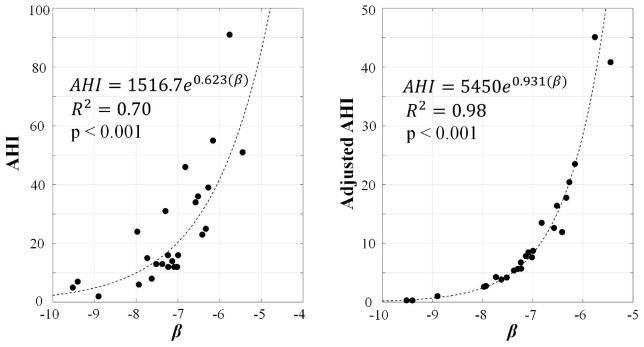

FIGURE 5. Exponential relationship between β and AHI (left panel) and adjusted AHI (right panel). The dotted lines represent the fitted relationship.

E. DETECTION ERROR

Paradoxical breathing is not exclusive to UA obstruction; it can occur independently of any pathology [46]. Furthermore, paradoxical breathing can be induced by the transition of sleep stages, particularly rapid eye movement [46], [47]. Moreover, other studies suggest that breathing becomes less asynchronous with age as the chest wall stiffens [48], [49]. To this end, no further analysis has been conducted on nonapneic paradoxical breathing. However, the less asynchronous breathing can be explained in the second stage, which also helps explain the elevated specificity. In addition to apnea annotation, the current database includes other annotations, such as *pb-event* (i.e., periodic event), when there is no proof of oxygen desaturation even when breathing has ceased for a period of 10 s or more. Other annotations, such as *possible*, is included for an event wherein a significant drop was observed, although cessation of breathing occurred in less than 10 s. Thus, it was observed that most *fp* were associated with these uncertainties. The remaining *fp*s can be attributed to other pathologies that induce paradoxical

breathing (i.e., sleep stages). On the one hand, for a therapeutic device, *fp* is almost always acceptable. For example, in the event of *fp*, the APAP machine will be false triggered, thereby delivering additional humidified air into the lungs. Thus, according to our understanding, *fp* may not cause severe implications to a patient. On the other hand, *fn* events are nontrivial and can cause false diagnostic decisions (i.e., a patient may be diagnosed as free of SAS when it is not be the case). For real-time applications, a significant count of *fn* events can results in a large number of missed interventions, which can lead to serious complications. Nevertheless, we have demonstrated that ORD can detect almost all severe obstructive apnea events that are not hypopnea. In addition, although *fn* can be observed, the ORD can predict AHI with high accuracy and help prevent false diagnostic decisions.

F. SIGNIFICANCE OF ORD

RIP signal quantification can act as surrogate estimation for other physiological conditions [8]. However, it may vary slightly depending on the signal processing of the RIP and whether the RIP calibration is considered. The calibrated and un-calibrated values of RIPsum provide an estimate of the tidal volume [9], [36], [37] and the relative change in tidal volume (compared to the baseline) [9], [20], [21], respectively. Furthermore, the calibrated and un-calibrated time derivatives of the RIPsum (RIPflow) provide semi-quantitative estimates of the airflow and relative airflow, respectively [9], [10], [23], [38], [39].

The key working mechanic of ORD is the dynamic measurement of the baseline condition upon every occurrence of *D*. Thus, the percentage change from baseline can be defined. Therefore, the relative change of ORD may be associated with the relative change in the tidal volume. This conclusion is based on the empirical evidence presented in previous studies. However, most clinical PSGs do not include numerical results associated with pneumotachograph recordings of respiratory gas flow to the lung. Hence, additional comparative analyses cannot be conducted.

The novelties of ORD include the introduction of an intermediate stage (S2) between nonapneic (S1) and apneic-related breathing (S3) and the appropriate definition of a threshold value (T_{ORD}). These implementations result in improved specificity ($p < 0.001$), as presented in Table 4. This result suggests that S2 can handle the effect of breathing inconsistencies, a certain degree of RIP belt distortion, and/or changes in body position during sleep. Furthermore, it was hypothesized that S2 can lead to additional qualitative changes in RIPsum. Fig. 4 presents the RIPsum, ORD, and BE sample-tracing for two different conditions. Depending on the circumstances, *D* can be regarded as an approximation of the OSA onset (or nadir), as manifested in the first OSA event. A slight difference is observed for the second OSA event, and further increase occurs in the following events. This largely results from a different rate of UA obstruction; *D* is defined when there are considerable excursions of the RIPsum amplitude. Furthermore, the RIPsum inspiratory portion changes are considerably more apparent several seconds before and after *D*, and this is observed for both conditions. This change may be associated with the inspiratory flow limitation (IFL). One study suggested that airflow limitations can be inferred from subtle qualitative changes in the RIP inspiratory portion [22].

According to the official American Thoracic Society report, a higher sampling frequency (25–50 Hz) is required for the qualitative analysis of the IFL pattern [40]. However, this is not the case in several RAAD studies, which adopted frequencies of 20 Hz [12] and 5 Hz [18]. Down-sampling was commonly performed when subtle changes in RIP were not of interest and/or to reduce processing speed. In this study, the original sampling frequency was maintained at 128 Hz, and it is a highly recommended value. Furthermore, per expert consensus, a key piece of evidence for IFL representation is the increase in the inspiratory duty cycle [41] and/or scooping for $\geq$75% of the inspiratory cycle duration [40]. This definition is consistent with RIPsum tracing (right panel), and it can be classified as an intermediate IFL (refer to the given pattern provided in [40]).

Manual visual IFL annotations require trained experts, which are very time consuming. Morover, the absence of IFL annotation in nearly all noninvasive PSG studies is a primary reason the quantification of subtle UA changes has been challenging. However, it is not our aim to perform the automatic detection of IFL. This should be left to the current automated IFL detection system [42]–[45]. Although the annotation for IFL was not provided with proper physiological reasoning (i.e., the S1$\rightarrow$S2$\rightarrow$S3 transition represents the progressive closure of the UA) and high sensitivity and specificity, our proposition that the region between the alpha and theta thresholds may represent a considerable reduction in the flow can be validated concomitantly.

V. FUTURE WORK AND CONCLUSION

This work demonstrated the state-of-the-art ORD approach. The advantages of the ORD algorithm are highlighted as follows:

- Finding a proper pre-baseline is difficult in automatic SAS detection [8], [20]. The ORD algorithm manages to ease the baseline paradox by defining a concise point (i.e., *D*) for a pre-baseline measurement.
- The ORD algorithm can distinguish obstructive apnea and hypopnea. This is possible when the pre-baseline is known; thus, an accurate drop in the RIPsum signal excursion can be measured. Furthermore, this algorithm can assist categorizing obstructive events in manual clinical diagnosis.
- This algorithm is suitable for real-time applications, because it is not computationally expensive.
- ORD is competent for further quantitative and qualitative IFL analysis. This finding will significantly benefit the automated IFL detection system studies.
- Owing to higher *SEN* and *SPE*, a novel parameter (β) can be defined from the ORD analysis to predict

exceptionally accurate AHI value regardless of patient demographics.

However, there are several limitations to ORD. The proposed approach cannot yet automatically distinguish OSA from hypopnea, because the three stages of transition-based classification restrain the current work scope. A manual distinction between OSA and hypopnea events was conducted to gain further insight into the ORD approach. However, this can be easily resolved. The peak distribution for a bin interval of 80–85% of $g_{\geq 4\%}$ distribution can be applied as additional information to distinguish between OSA and hypopnea precisely. As such, BE is now a four-stage expression.

In this study, CSA events were excluded because of the slight pathophysiology difference from obstructive events. CSA is characterized by the cessation of airflow with the absence of BE [50]. This is completely opposite to the mechanics of ORD, which predominantly exploits the paradoxical movement results from the increase in BE. When no BE is observed, continued excursion never occurs. Hence, the BE transition remains at either the first or the second stage. Differentiating between CSA and obstructive events using the BE has been a major concern to practitioners of the manual annotation process [8]. Failure to detect the BE can result in the misclassification of obstructive events, such as CSA [51]. Thus, although ORD cannot theoretically detect CSA, it can assist in the manual classification of obstructive events. However, further qualitative and quantitative studies on CSA using ORD need to be conducted.

The ORD algorithm was developed to be as generalized as possible. We considered a linearization function as denoted in Eq. (1). F and G are bounded parameters (i.e., normalized in the range of zero to one) that reduce the patient's variability. In addition, we considered inter-subject evaluation for more robust measurements. The patient demographics in the current database is quite dispersed (Table 1). Thus, the averaged value of T_{ORD} can be considered to be an optimal value and can be applied directly to other databases. However, it is unclear if we can achieve the same performance as ORD using different databases. It is possible to observe a slight change in the T_{ORD} value because of the different number of patients, their specific demographics (e.g., female only), or their specific diseases within other databases. Therefore, confirmatory studies should be conducted with other databases to re-determine ORD parameters, particularly T_{ORD}, before further generalization is conducted.

It is also compelling to conduct multivariate analysis using ORD with other physiological signals (i.e., SpO2, electroencephalogram, and heart-rate variability) to further understand the psychophysiological conditions of sleep apnea. For example, there are a few aspects of interest in the physiology behind heart-rate variability. Thus, it is interesting to investigate the time course of parasympathetic or sympathetic activation at different ORD stages. Parasympathetic activity may be prominent during the first stage. Meanwhile, an elevated sympathetic activity may be observed during the second and third stages. We intend to consider this in future work.

REFERENCES

[1] J. R. Tietjens, D. Claman, E. J. Kezirian, T. De Marco, A. Mirzayan, B. Sadroonri, A. N. Goldberg, C. Long, E. P. Gerstenfeld, and Y. Yeghiazarians, "Obstructive sleep apnea in cardiovascular disease: A review of the literature and proposed multidisciplinary clinical management strategy," *J. Amer. Heart Assoc.*, vol. 8, no. 1, Jan. 2019.

[2] R. C. Castello-Branco, T. Cerqueira-Silva, A. L. Andrade, B. M. M. Gonçalves, C. B. Pereira, I. F. Felix, L. S. B. Santos, L. M. Porto, M. E. L. Marques, M. B. Catto, M. A. Oliveira, P. R. S. P. de Sousa, P. J. R. Muiños, R. M. Maia, S. Schnitman, and J. Oliveira-Filho, "Association between risk of obstructive sleep apnea and cerebrovascular reactivity in stroke patients," *J. Amer. Heart Assoc.*, vol. 9, no. 6, pp. 1–9, Mar. 2020.

[3] J. Collen, C. Lettieri, E. Wickwire, and A. Holley, "Obstructive sleep apnea and cardiovascular disease, a story of ounders!" *Sleep Breath.*, vol. 24, no. 4, pp. 1299–1313, 2020.

[4] A. Culebras, "Cerebrovascular disease and the pathophysiology of obstructive sleep apnea," *Curr. Neurol. Neurosci. Rep.*, vol. 7, no. 2, pp. 173–179, 2007.

[5] M. Bergeron and S. L. Ishman, "Persistent obstructive sleep apnea burden on family finances and quality of life," *Otolaryngol. Neck Surg.*, vol. 7, Apr. 2021, Art. no. 0194599820986566.

[6] R. Ferber, "ASDA standards of practice: Portable recording in the assessment of obstructive sleep apnea," *Sleep*, vol. 17, no. 4, pp. 378–392, 1994.

[7] L. R. Kline, *Clinical Presentation and Diagnosis of Obstructive Sleep Apnea in Adults* (UpToDate), N. Collop, Ed. 2021.

[8] R. B. Berry, R. Budhiraja, D. J. Gottlieb, D. Gozal, C. Iber, V. K. Kapur, C. L. Marcus, R. Mehra, S. Parthasarathy, S. F. Quan, S. Redline, K. P. Strohl, S. L. D. Ward, and M. M. Tangredi, "Rules for scoring respiratory events in sleep: Update of the 2007 AASM manual for the scoring of sleep and associated events," *J. Clin. Sleep Med.*, vol. 8, no. 5, pp. 597–619, Oct. 2012.

[9] R. Farre, "Noninvasive monitoring of respiratory mechanics during sleep," *Eur. Respiratory J.*, vol. 24, no. 6, pp. 1052–1060, Dec. 2004.

[10] R. Thurnheer, X. Xie, and K. E. Bloch, "Accuracy of nasal cannula pressure recordings for assessment of ventilation during sleep," *Amer. J. Respiratory Crit. Care Med.*, vol. 164, no. 10, pp. 1914–1919, Nov. 2001.

[11] C. Coronel, "Comparison of PSG signals and respiratory movement signal via 3D camera in detecting sleep respiratory events by LSTM models," in *Proc. APSIPA ASC*, 2020, pp. 919–923.

[12] K. F. Whyte, M. Gugger, G. A. Gould, J. Molloy, P. K. Wraith, and N. J. Douglas, "Accuracy of respiratory inductive plethysmograph in measuring tidal volume during sleep," *J. Appl. Physiol.*, vol. 71, no. 5, pp. 1866–1871, Nov. 1991.

[13] M. A. Sackner, H. Watson, A. S. Belsito, D. Feinerman, M. Suarez, G. Gonzalez, F. Bizousky, and B. Krieger, "Calibration of respiratory inductive plethysmograph during natural breathing," *J. Appl. Physiol.*, vol. 66, no. 1, pp. 410–420, Jan. 1989.

[14] K. Montazeri, S. A. Jonsson, J. S. Agustsson, M. Serwatko, T. Gislason, and E. S. Arnardottir, "The design of RIP belts impacts the reliability and quality of the measured respiratory signals," *Sleep Breathing*, vol. 7, pp. 1–7, Jan. 2021.

[15] Y.-Y. Lin, H.-T. Wu, C.-A. Hsu, P.-C. Huang, Y.-H. Huang, and Y.-L. Lo, "Sleep apnea detection based on thoracic and abdominal movement signals of wearable piezoelectric bands," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 6, pp. 1533–1545, Nov. 2017.

[16] M. T. Bianchi, T. Lipoma, C. Darling, Y. Alameddine, and M. B. Westover, "Automated sleep apnea quantification based on respiratory movement," *Int. J. Med. Sci.*, vol. 11, no. 8, pp. 796–802, 2014.

[17] H. Azimi, S. S. Gilakjani, M. Bouchard, R. A. Goubran, and F. Knoefel, "Automatic apnea-hypopnea events detection using an alternative sensor," in *Proc. IEEE SAS*, Seoul, South Korea, Dec. 2018, pp. 1–5.

[18] T. Van Steenkiste, W. Groenendaal, D. Deschrijver, and T. Dhaene, "Automated sleep apnea detection in raw respiratory signals using long short-term memory neural networks," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 6, pp. 2354–2364, Nov. 2019.

[19] D. I. Loube, T. Andrada, and R. S. Howard, "Accuracy of respiratory inductive plethysmography for the diagnosis of upper airway resistance syndrome," *Chest*, vol. 115, no. 5, pp. 1333–1337, May 1999.

[20] American Academy of Sleep Medicine Task Force and others, "Sleep-related breathing disorders in adults: Recommendations for syndrome definition and measurement techniques in clinical research," *Sleep*, vol. 22, no. 4, pp. 667–689, 1999.

[21] G. A. Gould, "The sleep hypopnea syndrome," *Am. Rev. Respiratory Dis.*, vol. 137, no. 4, pp. 895–898, 1988.
[22] J. F. Masa, J. Corral, M. J. Martín, J. A. Riesco, A. Sojo, M. Hernández, and N. J. Douglas, "Assessment of thoracoabdominal bands to detect respiratory effort-related arousal," *Eur. Respiratory J.*, vol. 22, no. 4, pp. 661–667, Oct. 2003.
[23] V. Kaplan, J. N. Zhang, E. W. Russi, and K. E. Bloch, "Detection of inspiratory flow limitation during sleep by computer assisted respiratory inductive plethysmography," *Eur. Respiratory J.*, vol. 15, no. 3, pp. 570–578, 2000.
[24] A. M. O'Mahony, J. F. Garvey, and W. T. McNicholas, "Technologic advances in the assessment and management of obstructive sleep apnoea beyond the apnoea-hypopnoea index: A narrative review," *J. Thoracic Disease*, vol. 12, no. 9, pp. 5020–5038, Sep. 2020.
[25] *St. Vincent's University Hospital/University College Dublin Sleep Apnea Database*. Accessed: Feb. 1, 2020. [Online]. Available: https://archive.physionet.org/pn3/ucddb/
[26] H. Azimi, P. Xi, M. Bouchard, R. Goubran, and F. Knoefel, "Machine learning-based automatic detection of central sleep apnea events from a pressure sensitive mat," *IEEE Access*, vol. 8, pp. 173428–173439, 2020.
[27] A. V. Vroegop, O. M. Vanderveken, A. N. Boudewyns, J. Scholman, V. Saldien, K. Wouters, M. J. Braem, P. H. Van de Heyning, and E. Hamans, "Drug-induced sleep endoscopy in sleep-disordered breathing: Report on 1,249 cases," *Laryngoscope*, vol. 124, no. 3, pp. 797–802, Mar. 2014.
[28] N. Jain, J. Rodin, M. S. Boon, and C. T. Huntley, "A systematic approach to the evaluation and management of obstructive sleep apnea: The jefferson protocol," *Amer. J. Otolaryngol.*, vol. 42, no. 2, Mar. 2021, Art. no. 102866.
[29] C. Coronel, C. Wiesmeyr, H. Garn, B. Kohn, M. Wimmer, M. Mandl, M. Glos, T. Penzel, G. Klösch, A. Stefanic-Kejik, M. Böck, E. Kaniusas, and S. Seidel, "Detection of respiratory events by respiratory effort and oxygen desaturation," *J. Med. Biol. Eng.*, vol. 40, no. 4, pp. 517–525, Aug. 2020.
[30] N. M. Punjabi, A. B. Newman, T. B. Young, H. E. Resnick, and M. H. Sanders, "Sleep-disordered breathing and cardiovascular disorders: An outcome-based definition of hypopneas," *Respiratory Care*, vol. 177, pp. 1150–1155, 2008.
[31] A. Sabil, M. Le Vaillant, C. Stitt, F. Goupil, T. Pigeanne, L. Leclair-Visonneau, P. Masson, A. Bizieux-Thaminy, M.-P. Humeau, N. Meslier, and F. Gagnadoux, "A CPAP data–based algorithm for automatic early prediction of therapy adherence," *Sleep Breathing*, pp. 1–6, Sep. 2020.
[32] H. Fox, T. Bitter, O. Sauzet, V. Rudolph, and O. Oldenburg, "Automatic positive airway pressure for obstructive sleep apnea in heart failure with reduced ejection fraction," *Clin. Res. Cardiol.*, vol. 7, pp. 1–10, Jul. 2020.
[33] Mostafa, Mendonça, Ravelo-García, and Morgado-Dias, "A systematic review of detecting sleep apnea using deep learning," *Sensors*, vol. 19, no. 22, p. 4934, Nov. 2019.
[34] J. A. Waxman, D. Graupe, and D. W. Carley, "Automated prediction of apnea and hypopnea, using a LAMSTAR artificial neural network," *Amer. J. Respiratory Crit. Care Med.*, vol. 181, no. 7, pp. 727–733, Apr. 2010.
[35] B. L. Koley and D. Dey, "Real-time adaptive apnea and hypopnea event detection methodology for portable sleep apnea monitoring devices," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 12, pp. 3354–3363, Dec. 2013.
[36] M. J. Tobin, "Breathing abnormalities during sleep," *Arch. Internal Med.*, vol. 143, no. 6, pp. 1221–1228, Jun. 1983.
[37] J. P. Cantineau, P. Escourrou, R. Sartene, C. Gaultier, and M. Goldman, "Accuracy of respiratory inductive plethysmography during wakefulness and sleep in patients with obstructive sleep apnea," *Chest*, vol. 102, no. 4, pp. 1145–1151, Oct. 1992.
[38] S. A. Clark, C. R. Wilson, M. Satoh, D. Pegelow, and J. A. Dempsey, "Assessment of inspiratory flow limitation invasively and noninvasively during sleep," *Amer. J. Respiratory Crit. Care Med.*, vol. 158, no. 3, pp. 713–722, Sep. 1998.
[39] A. Griffiths, J. Maul, A. Wilson, and S. Stick, "Improved detection of obstructive events in childhood sleep apnoea with the use of the nasal cannula and the differentiated sum signal," *J. Sleep Res.*, vol. 14, no. 4, pp. 431–436, Dec. 2005.
[40] S. Pamidi, "An official American thoracic society workshop report: Noninvasive identification of inspiratory flow limitation in sleep studies," *Ann. Am. Thoracic Soc.*, vol. 14, no. 7, pp. 1076–1085, 2017.
[41] H. Schneider, V. Krishnan, L. E. Pichard, S. P. Patil, P. L. Smith, and A. R. Schwartz, "Inspiratory duty cycle responses to flow limitation predict nocturnal hypoventilation," *Eur. Respiratory J.*, vol. 33, no. 5, pp. 1068–1076, May 2009.
[42] R. G. Norman, D. M. Rapoport, and I. Ayappa, "Detection of flow limitation in obstructive sleep apnea with an artificial neural network," *Physiol. Meas.*, vol. 28, no. 9, pp. 1089–1100, Sep. 2007.
[43] T. Aittokallio, T. Saaresranta, P. Polo-Kantola, O. Nevalainen, and O. Polo, "Analysis of inspiratory flow shapes in patients with partial upper-airway obstruction during sleep," *Chest*, vol. 119, no. 1, pp. 37–44, Jan. 2001.
[44] C. Morgenstern, M. Schwaibold, W. Randerath, A. Bolz, and R. Janeá, "Automatic non-invasive differentiation of obstructive and central hypopneas with nasal airflow compared to esophageal pressure," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.*, Aug. 2010, pp. 6142–6145.
[45] H. Takamoto, H. Nishine, S. Sato, G. Sun, S. Watanabe, K. Seokjin, M. Asai, M. Mineshita, and T. Matsui, "Development and clinical application of a novel non-contact early airflow limitation screening system using an infrared time-of-flight depth image sensor," *Frontiers Physiol.*, vol. 11, pp. 1–9, Sep. 2020.
[46] J. Hammer and C. J. L. Newth, "Assessment of thoraco-abdominal asynchrony," *Paediatric Respiratory Rev.*, vol. 10, no. 2, pp. 75–80, Jun. 2009.
[47] M. D. Goldman, M. Pagani, H. T. T. Trang, J.-P. Praud, R. Sartene, and C. Gaultier, "Asynchronous chest wall movements during non-rapid eye movement and rapid eye movement sleep in children with bronchopulmonary dysplasia," *Amer. Rev. Respiratory Disease*, vol. 147, no. 5, pp. 1175–1184, May 1993.
[48] J. Kohyama, "Asynchronous breathing during sleep," *Arch. Disease Childhood*, vol. 84, no. 2, pp. 174–177, Feb. 2001.
[49] M. B. Hershenson, A. A. Colin, M. E. B. Wohl, and A. R. Stark, "Changes in the contribution of the rib cage to tidal breathing during infancy," *Amer. Rev. Respiratory Disease*, vol. 141, no. 4, pp. 922–925, Apr. 1990.
[50] D. J. Eckert, A. S. Jordan, P. Merchia, and A. Malhotra, "Central sleep apnea: Pathophysiology and treatment danny," *Chest*, vol. 131, no. 2, pp. 595–607, 2007.
[51] A. Boudewyns, M. Willemen, M. Wagemans, W. De Cock, P. Van de Heyning, and W. De Backer, "Assessment of respiratory effort by means of strain gauges and esophageal pressure swings: A comparative study," *Sleep*, vol. 20, no. 2, pp. 168–170, Feb. 1997.

MUHAMMAD SHAUFIL ADHA received the B.S. and M.S. degrees in electrical and electronics engineering from the University of Technology Malaysia, Johor Bahru, Malaysia, in 2014 and 2017, respectively. He is currently pursuing the Ph.D. degree with the Graduate School of Science and Technology, Kumamoto University, Kumamoto, Japan.

His research interest includes applied signal processing with an emphasis on physiological signal processing.

TOMOHIKO IGASAKI received the B.S., M.S., and Ph.D. degrees in engineering from Kumamoto University (KU), Kumamoto, Japan, in 1994, 1996, and 2007, respectively.

In 1999, he joined the Faculty of Engineering, KU, as a Research Associate. He was appointed as an Assistant Professor with the Graduate School of Science and Technology (GSST), KU, in 2007. Since 2009, he has been an Associate Professor with GSST (reorganized into the Faculty of Advanced Science and Technology, in 2016), KU. Since 2017, he also serves as a Special Assistant to the President of KU for international student employment support. His research interest includes biological signal processing (quantification/visualization for practical use, especially health/medical care). He is a Councilor of the Japanese Association for the Study of Taste and Smell, a Delegate of the Japanese Society of Clinical Neurophysiology and the Japanese Society of Medical and Biological Engineering (a trustee of Kyushu branch), a member of the Institute of Electronics, Information, and Communication Engineers, and an affiliate of the IEEE Engineering in Medicine and Biology Society.