Contactless Sleep Apnea Detection on Smartphones

Rajalakshmi Nandakumar
Computer Science and Engineering
University of Washington
rajaln@uw.edu

Shyamnath Gollakota
Computer Science and Engineering
University of Washington
gshyam@uw.edu

Nathaniel Watson M.D.
UW Medicine Sleep Center
University of Washington
nwatson@uw.edu

Abstract – We present a contactless solution for detecting sleep apnea events on smartphones. To achieve this, we introduce a novel system that monitors the minute chest and abdomen movements caused by breathing on smartphones. Our system works with the phone away from the subject and can simultaneously identify and track the fine-grained breathing movements from multiple subjects. We do this by transforming the phone into an active sonar system that emits frequency-modulated sound signals and listens to their reflections; our design monitors the minute changes to these reflections to extract the chest movements. Results from a home bedroom environment shows that our design operates efficiently at distances of up to a meter and works even with the subject under a blanket.

Building on the above system, we develop algorithms that identify various sleep apnea events including obstructive apnea, central apnea, and hypopnea from the sonar reflections. We deploy our system at the UW Medicine Sleep Center at Harborview and perform a clinical study with 37 patients for a total of 296 hours. Our study demonstrates that the number of respiratory events identified by our system is highly correlated with the ground truth and has a correlation coefficient of 0.9957, 0.9860, and 0.9533 for central apnea, obstructive apnea and hypopnea respectively. Furthermore, the average error in computing of rate of apnea and hypopnea events is as low as 1.9 events/hr.

CATEGORIES AND SUBJECT DESCRIPTORS

J.3 [**Computer Applications**]: Life and Medical Sciences

GENERAL TERMS

Design; Human Factors; Algorithms

KEYWORDS

Mobile Health; Sleep Apnea; Phone Sonar; Contactless Breathing Monitoring

1. INTRODUCTION

Sleep apnea is a common medical disorder that occurs when breathing is disrupted during sleep. It is estimated to affect more than 18 million American adults [9, 41] and is linked to attention deficit/hyperactivity disorder, high blood pressure, diabetes, heart
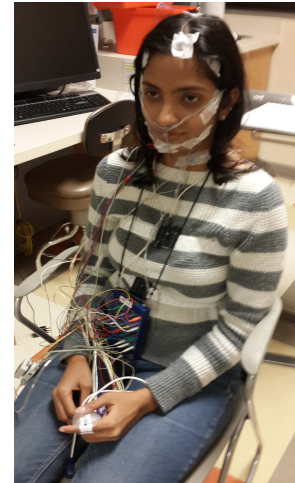
Figure 1—Sensors used in the polysomnography test. The figure shows all the sensors used in the test along with the data collection unit. Polysomnography is used to diagnose various sleep disorders including sleep apnea. Our goal is to use a smartphone to detect sleep apnea events without any sensors on the human body.

attack, stroke, and increased motor vehicle accidents [6, 39]. Diagnosing sleep apnea in the clinic requires the polysomnography test which is an expensive, time-consuming and labor-intensive process. It requires a trained technician to attach and monitor various sensors on the patient for the sleep duration and is typically associated with long waiting lists [21]. While portable recording systems are being developed for use in home settings, they require instrumenting either the patient [25, 20, 20] or the bed [30] with various sensors and most still require a trained technician to setup the recording system [30].

In this paper we ask the following question: Can we leverage smartphones to detect sleep apnea events without the need for sensor instrumentation? The key challenge is that detecting sleep apnea events requires tracking the fine-grained abdomen and chest movements due to breathing [18]. While the iPhone Respiratory app [4] can track the breathing movements, it requires placing the phone on the body between the ribcage and the stomach and hence is intrusive. Vision-based solutions [43] can track these movements without instrumenting users, but are limited to line-of-sight and good lighting conditions and hence are not applicable to the sleep environment, i.e., in the dark or under a blanket.

We introduce a novel contactless system that tracks the chest and abdomen movements on smartphones and works in the sleep environment. It operates with the phone away from the user and can concurrently track the breathing movements from multiple users. Using this design, we build *ApneaApp*, a smartphone-based solution for detecting sleep-related respiratory events reported in a clinical polysomnography test including hypopnea (when the subject's

breathing becomes shallow), obstructive apnea (a complete or partial blockage of the subject's airway) and central apnea (when the subject holds his or her breath).

Our key insight is to transform the phone into an active sonar system. At a high level, we transmit 18-20 kHz sound waves from the phone speaker and listen to their reflections at the microphone. The chest and abdomen motion due to breathing creates changes to the reflected sound waves. These changes, however, are minute and extracting them reliably from other environmental reflections is challenging. To overcome this, we employ FMCW (frequency modulated continuous wave) transmissions that allow us to separate reflections arriving at different times by mapping time differences to shifts in the carrier frequency. Specifically, the reflections from the human body arrive at a specific time depending on the distance from the phone speaker. Thus, focusing on the corresponding frequency allows us to reliably extract the amplitude changes due to breathing, in the presence of all other environmental reflections. Further, since reflections from multiple subjects would arrive at different times, the corresponding frequencies provide us with the ability to simultaneously track multiple breathing signals. Finally, non-breathing body motion creates reflection patterns distinct from breathing, enabling us to distinguish between them.

We implement our design on off-the-shelf smartphones and run benchmark experiments with five healthy participants in a bedroom environment using the Vernier respiratory belt as a baseline. Our results show the following:

- Our system estimates the coarse-grained breathing frequency[1] to within 99.2% of the baseline at distances of up to a meter from the subject. This translated to an error of less than 0.11 breaths/min. These accuracies remain this high even when the subjects use blankets.
- The above accuracies remain unaffected by audible noise in the environment from vehicles on a nearby street as well as human conversations. This is because, we use a high-pass filter to filter out audible signals below 18 kHz.
- It can separate and concurrently track the breathing movements of two subjects on the bed separated by 20 cm.

Building on the above system, we design algorithms to compute the number of central, obstructive and hypopnea events as well as the apnea-hypopnea index which is the average rate of apnea and hypopnea events during the sleep duration. We achieve this by processing both the fine- and coarse- grained changes due to the chest and abdomen movements as well as non-breathing body motion. We deploy ApneaApp at the UW Medicine Sleep Center at Harborview and perform a clinical study with 37 patients for a total of 296 hours. The patients in our study were ordered by their physicians to undergo the polysomnography (PSG) test. Our study was done concurrently with the PSG test and we consider the sensor data and diagnosis from the latter as the ground truth for evaluating our system. Our study shows the following:

- Across patients, the number of central apnea, hypopnea and obstructive apnea events detected by our system is highly correlated with the ground truth. Specifically, the intra-class correlation coefficient between PSG and ApneaApp, is 0.9957, 0.9533 and 0.9860 for central apnea, hypopnea and obstructive apnea respectively.

---

[1]Detecting sleep apnea events requires tracking the fine-grained abdomen and chest motion variations in addition to the coarse-grained breathing frequency. We evaluate ApneaApp's ability to track these variations in our clinical study.

Respiratory Summary:
Types of Respiratory Events / Respiratory Effort Related Arousal (RERA) Events

Respiratory Events	Number	Index	RERA	Parameter	Total	Index
Obstructive Apneas	123	15.7 /hr		Total:	129	16.4
Mixed Apneas	0	0.0 /hr		Non-REM:	105	16.5
Central Apneas	1	0.1 /hr		REM:	24	16.1
Total Apneas	**124**	**15.8 /hr**		Supine:	129	16.4
Total Hypopneas*	**371**	**47.2 /hr**		Lateral:	N/A	N/A
Apneas + Hypopneas*	**495**	**63.0 /hr**		Prone:	N/A	N/A

Figure 2—Snapshot of a Clinical PSG Report. It summarizes the number of obstructive, central and hypopnea events along with the apneas-hypopneas index (AHI). An AHI value between 0–5 is classified as no-apnea, values between 5–15 are classified as mild-apnea, AHI values between 15–30 are classified as moderate-apnea, and higher AHIs are severe apnea conditions.

- The average error in computing the rate of apnea and hypopnea events is 1.9 events/hr; this is a clinically acceptable value [30].
- Our system accurately classifies 32 out of 37 patients between four sleep apnea levels (no apnea, mild, moderate, and severe apnea). The five misclassifications occur between no-apnea and mild-apnea; four of them happen right at the boundary between the two levels with an error less than 1 event/hr. These boundary cases are handled separately by physicians depending on the patient preferences, symptoms, and insurance; thus, effectively reducing the number of misclassifications to one.
- We ran an audibility test with 87 sleep apnea patients (ages between 23 and 93 with a mean age of 50) and 57 healthy undergraduate students at UW CSE. None of the 87 sleep apnea patients reported any audible sounds from ApneaApp. Only two of the 57 undergraduates reported hearing audible sounds. This demonstrates that ApneaApp is inaudible for most of the adult population.

Contributions: We make four key contributions: (1) We introduce a novel contactless technique for tracking chest and abdomen movements due to breathing on smartphones. We achieve this by analyzing the reflections from FMCW sonar transmissions. (2) We design algorithms to detect central apnea, obstructive apnea, and hypopnea as well as estimate the apnea-hypopnea index from the sonar reflections. (3) We implement our design on off-the-shelf smartphones and demonstrate the ability to concurrently track breathing movements from multiple subjects. (4) We perform a clinical study with 37 patients demonstrating the feasibility of using our system to accurately compute the number of central, obstructive, and hypopnea events as well as the apnea-hypopnea index.

2. POLYSOMNOGRAPHY OVERVIEW

The clinical polysomnography test (PSG) is traditionally used to diagnose sleep apnea and other sleep disorders. It is conducted overnight in a sleep laboratory where a trained technician monitors the patient's sleeping patterns. To do this, the technician attaches the patient with a number of sensors including a chest and abdomen belt to measure breathing movements, a nasal pressure transducer, a snore microphone, a pulse oximeter to measure oxygen saturation, a movement sensor on each leg to detect movements and five EEG sensors to measure brain activity. The sensors are all connected using wires and the technician monitors the live data stream from the sensors, throughout the sleep duration.

Fig. 2 shows a snapshot of a PSG report. The key metric used for sleep apnea diagnosis is the AHI — the Apnea-Hypopnea Index — that represents the rate at which apnea and hypopnea events occur during the sleep period. Physicians classify the sleep apnea level using these AHI values. Specifically, AHI values between 0–5 are classified as no-apnea, those between 5–15 are classified as mild-

(a) Central Apnea Event

(b) Hypopnea Event

(c) Obstructive Apnea Event
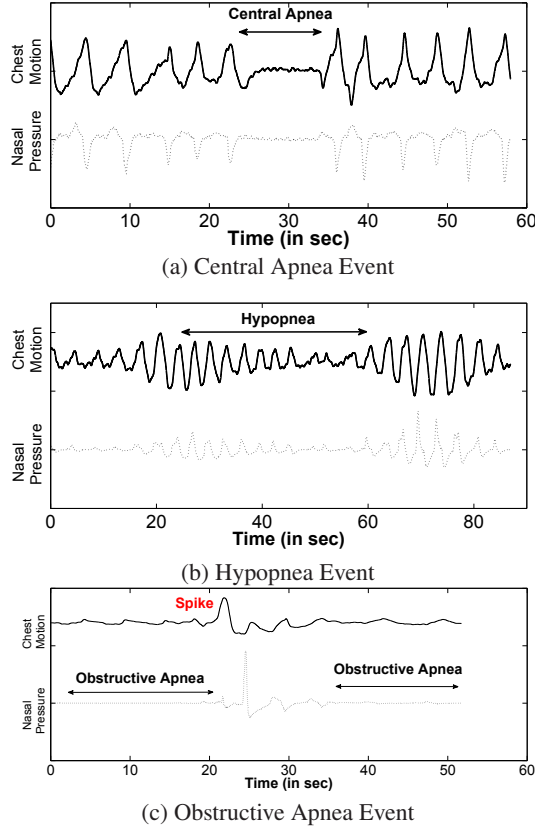
Figure 3—American Academy of Sleep Medicine (AASM) Signal Characterization of the Apnea Events. The figures show the chest motion and nasal pressure signals for the three apneas. A central apnea event occurs when the subject holds her breath for a non-negligible duration. A hypopnea event occurs when the subject's chest motion drops by more than 30% with an accompanying 3–4% oxygen desaturation. Finally, an obstructive apnea event occurs when the subject makes an increased effort to pull air into the lungs but air does not reach the lungs due to blockage.

apnea, AHI values between 15–30 are classified as moderate-apnea, and higher AHIs are severe apnea.

The apnea-hypopnea index is computed as follows:

$$AHI = \frac{\#central\ apnea + \#hypopnea + \#obstructive\ apnea}{total\ sleep\ time}$$

In the above equation, central apnea, hypopnea, and obstructive apnea denote the various apnea conditions that are tracked during the study. Mixed apneas are another class of apneas that are sometimes included in the above equation. However, none of our PSG reports showed non-zero mixed apneas and so we ignore them in our computation.

To compute the above parameters, the eight-hour sensor data is split into 30-second intervals called epochs. The scoring process of analyzing these epochs involves two main steps. The first step is staging, which identifies whether the patient is awake or asleep in each epoch. This is achieved by examining the brain activity obtained from the five EEG sensors. At the end of this step, each epoch is marked as being in either a wake or sleep stage. The second step involves identifying the number of central apnea, hypopnea, and obstructive apnea events, using the AASM guidelines [18] outlined below.

Identifying central apnea events. A central apnea event occurs when the subject holds her breath for a non-negligible duration. Fig. 3(a) shows the nasal pressure and chest motion signals during a central apnea event. The figure shows that the chest movements are flat indicating the absence of breathing effort; consequentially the nasal pressure is also flat. If this persists for more than ten seconds, it is marked as a central apnea event.

Identifying hypopnea events. A hypopnea event occurs when the subject's breathing becomes shallow. Fig. 3(b) plots the nasal pressure and chest motion signals during a hypopnea event. The figure shows that during a hypopnea event, the chest movements reduce in amplitude. In particular, if this amplitude drops by more than 30% and has an accompanying 3–4% oxygen desaturation, then the corresponding epoch is marked as hypopnea. We note that recent clinical research [30] has shown that the 30% reduction alone can be used for detecting hypopneas without a significant loss of accuracies.

Identifying obstructive apnea events. Obstructive apnea occurs when there is a complete or partial blockage of the upper airway during sleep. During an obstructive apnea event, the subject makes an effort to pull air into the lungs, however air does not reach the lungs because of blockage. Fig. 3(c) shows the signals where the breathing effort can be seen in the chest band signals and the air flow is flat in the nasal pressure sensor.

We note the following three points about PSG:

- The current procedure for sensor data collection and processing is both labor and time intensive. Specifically, it takes about an hour for the technician to fit each patient with these sensors. Throughout a sleep duration of eight hours, the technician monitors the sensors and ensures that they remain properly attached to the patient's body. The sensor data is then processed manually to tag every epoch with the sleep apnea events.
- While portable sleep apnea testing is performed in the home, it still require setting up the patient with chest and abdomen belts, nasal pressure sensors, transducer and thermistors, EKG and pulse oximetry. Home testing has a high failure rate of up to 33% due to the loss of signal resulting from detachment of wires and cables [33].
- A PSG test is also used to diagnosis other sleep-related conditions including upper airway resistance syndrome which involve respiratory effort related arousals (RERA) that are shown in Fig. 2. RERAs are sleep arousals that do not meet the above definitions of apneas and hypopneas. While these are respiratory-related and could be detected using our sonar-based system, exploring them in detail is not in the scope of this paper.

3. APNEAAPP

ApneaApp is a contactless system that enables detection of sleep apnea events using smartphones. To understand how ApneaApp operates, we first describe how we transform the phone into an active sonar system that tracks the chest and abdomen movements due to breathing. We then describe our algorithms to detect sleep apnea events from these movements.

3.1 Transforming the Phone into an Active Sonar

An FMCW waveform is a chirp signal, as shown in Fig. 4, where the transmitted frequency increases linearly with time between 18 kHz and 20 kHz. These signals reflect off the reflector (e.g., human body) and arrive at the microphone after a time delay. To determine this delay, at a high level, an FMCW receiver com-
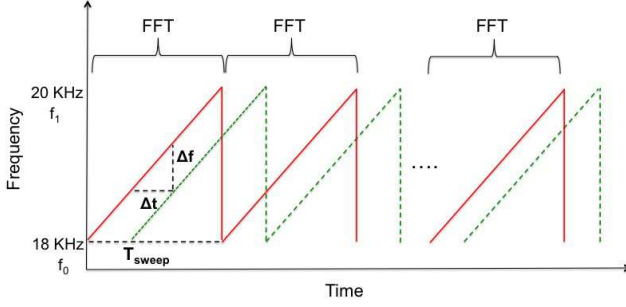
Figure 4—Traditional FMCW Processing. The transmitter continuously transmits signals where the frequency increases linearly with time between f_0 and f_1. The reflections that arrive with a time delay Δt create a frequency shift Δf. The receiver extracts this frequency shift by performing an FFT over the chirp duration.

pares the frequencies of the transmitted and reflected signals. Since the transmitted frequency increases linearly in time, time delays in the reflected signals translate to frequency shifts in comparison to the transmitted signals.

For instance, the red line is the transmitted signal from the phone speaker and the green line is the reflected signal from a human body that arrives with a time delay Δt. This delay is given by $\frac{2d}{v_{sound}}$, where d is the distance from the human body and v_{sound} is the speed of sound. Now, the frequency shift Δf between the transmitted and reflected signal is:

$$\Delta f = \frac{f_1 - f_0}{T_{sweep}} \Delta t$$

When we have multiple reflectors that are at different distances from the receiver, their reflections translate to different frequency shifts in the signal. An FMCW receiver can extract all these frequency shifts by performing an Fourier transform over a chirp duration as shown in Fig. 4. The chirp duration, T_{sweep}, in practice is picked so that the reflections from all points within the desired operational distance would start arriving before the chirp ends. Since our operational distance is a meter we pick a chirp duration of 10.75 ms in our implementation.

Challenge: The act of breathing creates minute chest and abdomen motion that can be captured by monitoring the corresponding bin in the Fourier transform as a function of time. The challenge however is that breathing movements are minute and create a very small frequency shift. Specifically, a 2 cm breathing displacement creates a 11.7 Hz shift.[2] This is problematic because given our sampling rate and chirp duration, the width of each FFT bin is 93.75 Hz which is much greater than the frequency shifts created due to breathing.

To address this problem, as shown in Fig. 5, the ApneaApp receiver performs a Fourier transform over N FMCW chirps. This is in contrast to a traditional FMCW receiver that computes a Fourier transform over the duration of a single FMCW chirp. Such an operation, decreases the width of each FFT bin by a factor of N. In our implementation we set N to ten which results in an FFT bin width of 9.37 Hz. This allows us to capture the minute frequency shifts resulting from the breathing movements. We note that performing an

---

[2]Given the speed of sound, a 48 kHz sampling rate translates to a resolution of 0.71 cm per sample. Further, a 10.7 ms chirp duration corresponds to 512 samples. With 18–20 kHz FMCW chirps, each sample corresponds to a 3.9 Hz frequency shift. Thus, a displacement of 0.71 cm translates to a 3.9 Hz change in the frequency domain. Consequentially, a 2 cm breathing movement creates a 11.7 Hz frequency shift.
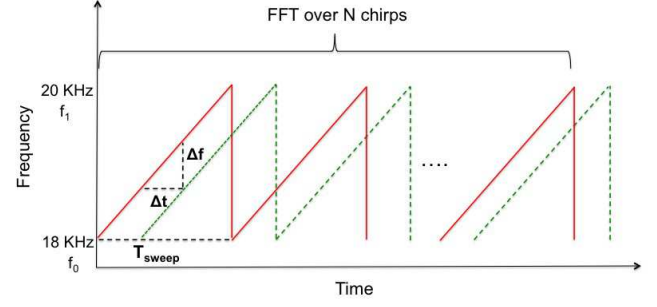
Figure 5—FMCW Processing in ApneaApp. To extract the minute frequency shifts created by breathing motion, ApneaApp performs an FFT over an integer number of chirp durations.

FFT over multiple chirp durations reduces our ability to track high-frequency movements that occur during these chirps. However, in our implementation, ten chirps correspond to a very short 107 ms, a duration within which significant breathing movements are unlikely to occur.

The final question is: *how do we compute the distance of the subject from the phone?* At a high level, we start at a distance of zero and search for breathing movements at increasing distance values up to the maximum distance of one meter. Specifically, we search for breathing movements in the 58 Fourier bin corresponding to 18 kHz to 18.546 kHz.[3] In our implementation we reduce the computation by searching in every alternate FFT bin. To search for these breathing movements in each FFT bin, we perform another FFT over a 30s duration and search for peaks in the typical breathing frequencies of 0.2-0.3 Hz.

To summarize, the phone transmits FMCW signals in the 18-20 kHz range with a chirp duration of 10.7 ms from its speaker. The microphones receive the reflected signals and process them to track the breathing movements. Specifically, we first find the distance to the human by searching for a periodic breathing signal starting from the closed distance value to the maximum range of one meter. Once we find this distance, we track the breathing movements by performing in a shorter FFT over ten chirp durations and monitor the reflected signals corresponding to the estimated distance value. Note that the above procedure is repeated every time the user moves their position, which we identify using the algorithm in §3.2.2. This prevents the need for manually calibrating the distance between the user and the phone.

We note the following points about our algorithm.

Computational Complexity. Our algorithm requires one 5120-point FFT and between one to twenty nine 24000-point FFTs to successfully extract the breathing motion. We stop our search at the first FFT bin that has the breathing movements.

Tracking breathing from multiple subjects. Reflections corresponding to subjects at different distances arrive with different time delays. Thus they create different frequency shifts with the FMCW signal. Therefore, to track breathing from two subjects, we modify the above algorithm to continue its search until it finds two FFT bins with the breathing motion.

Leveraging angle-of-arrival algorithms. One could track breathing movements from equidistant subjects who are at different angles, by using multiple microphones and implementing angle-of-arrival

---

[3]For 18-20 kHz FMCW transmissions, a distance of zero corresponds to the FFT bin for 18-kHz. The operational distance of 1 m corresponds to a frequency shift of 18.546 Hz.

algorithms. Evaluating this, however, is not in the scope of this paper.

FMCW versus pulse-modulated transmissions. Pulse-modulated transmissions use high-amplitude short pulses and are an alternative to continue-wave FMCW signals. In our experiment, however, they created low frequency components in the 0-18 kHz range that made them noticeably audible. FMCW transmissions, on the other hand, have lower-amplitudes and are limited to 18-20 kHz, making them inaudible for most of the adult population.

3.2 Sleep Apnea Detection Algorithm

As described in §2, diagnosing sleep apnea requires estimating the Apnea-Hypopnea Index (AHI) which is the average rate of apnea events during the sleep duration. This requires computing the number of central, obstructive, and hypopneas as well the total sleep time. In this section, we first describe our algorithms to compute the number of apneas and then the total sleep time.

3.2.1 *Estimating the Number of Apneas*

ApneaApp detects a hypopnea event when the chest motion reduces below a threshold (30%).[4] A central apnea event is detected when the subject holds her breath and as a result the amplitude of the chest motion signal reduces to zero. Thus, ApneaApp identifies hypopnea when the amplitude of the breathing motion decreases below the threshold but the periodicity that is expected in a breathing signal still exists. But to identify central apnea, we couple an amplitude reduction in the chest motion signal with an absence of the breathing periodicity.

An obstructive apnea event, on the other hand, occurs when there is an obstruction in the airflow, i.e., the subject makes a breathing effort but the airflow is obstructed by a tissue overgrowth in the neck. In a clinical PSG study, obstructive apnea is detected using the nasal pressure sensor that directly measures the airflow; this, however, is not available in our system. We instead perform an analysis of the chest motion data that reveals that the subject usually tends to increase her breathing effort in these scenarios. This results in a clear spike in the amplitude of the chest motion, as shown in Fig. 3(c).

Thus, measuring the amplitude and periodicity in the chest motion signal is critical to detecting obstructive, central and hypopneas. Note that as shown in Fig. 3, the chest motion signal can be approximated as a periodic sinusoidal wave. Hence the magnitude of the peaks of these sinusoidal waves represents the amplitude and their peak locations determine their periodicity. Thus, we design a peak detection algorithm to compute the amplitude and periodicity of the chest motion signal.

ApneaApp's peak detection algorithm. Standard peak detection algorithms identify the transition point at which the signal changes from an increasing to a decreasing trend. In other words, for every set of three points, if the middle point is the maximum, then it is labeled as a peak. Such an algorithm, however, would result in a number of erroneous peaks with our chest motion signal. As an example, Fig. 6 plots a typical chest motion signal. Running the standard peak detection algorithm on this signal results in a number of unintended peaks as shown in the figure. To reduce the number of such peaks, we introduce two key heuristics.

---

[4]AASM and Medicare guidelines require the 30% reduction to be accompanied with a 3% and 4% oxygen saturation respectively. However, recent clinical research [30] has shown that the 30% reduction alone can be used for detecting hypopneas.
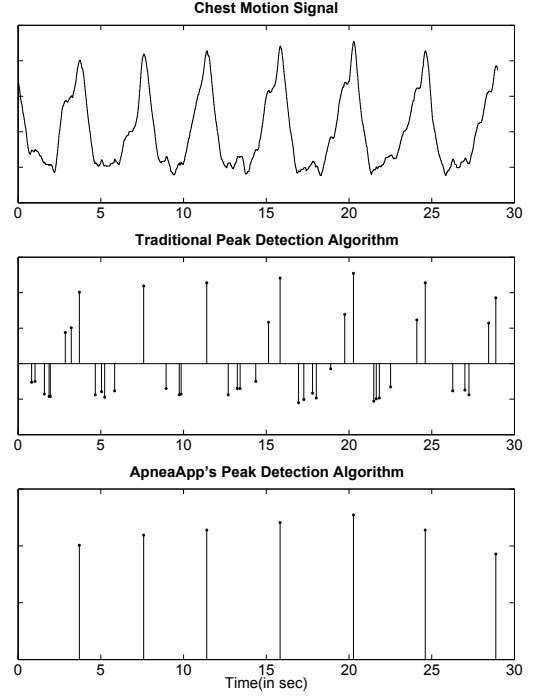

Figure 6—Understanding ApneaApp's peak detection algorithm. Standard peak detection algorithms identify the transition point at which the signal changes from an increasing trend to a decreasing trend. Thus, they identify a number of unintended peaks in our chest motion signal. ApneaApp's peak detection algorithm, in contrast, uses minimum peak distance and amplitude heuristics to reduce these unintended peaks.

The first heuristic is to set a threshold on the minimum distance between two consecutive peaks. In particular, the breathing frequency in an adult human typically varies between 12-18 breaths/min. Thus each breath takes 3.3 s at the maximum frequency of 18 breaths/min. We set a conservative threshold of three seconds in our implementation.

The second heuristic is to set a threshold on the minimum amplitude at which a peak is detected. To do this, we first run the above peak detection algorithm with the minimum distance heuristic on the chest motion data for the first hour to obtain an initial set of peaks. We then compute the minimum amplitude threshold as $\mu_{peaks} - 2\sigma_{peaks}$, where μ_{peaks} and σ_{peaks} are the mean and standard deviation of the peak amplitudes. Finally, we go back and apply both the amplitude as well as the minimum distance thresholds on the entire chest motion data for the eight-hour sleep duration to obtain the actual set of peaks. Fig. 6 shows that our peak detection algorithm identifies the correct peaks.

Central apnea estimation algorithm. We run the peak detection algorithm to identify the locations of the peaks in the chest motion signal. We then compute the distance between these consecutive peaks. If this distance is greater than ten seconds, it means that the subject holds her breath for a non-negligible period of time and hence we declare it as a central apnea event.

Hypopnea estimation algorithm. We again use the peak detection algorithm to detect the peaks. When the peak values reduce beyond a threshold and still maintain their periodicity, we declare it as a hypopnea event. To compute this threshold for our sonar data, we perform a linear regression on the data from a single patient to max-
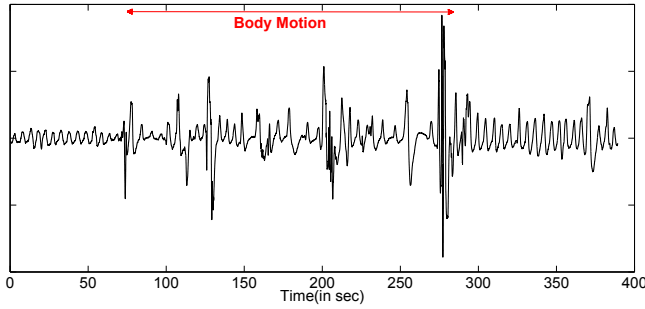
Figure 7—Differentiating body movements from breathing. ApneaApp leverages that when the subject moves her body, the sonar reflections experience large variations that do not have the periodicity of the breathing motion.

imize the hypopnea detection accuracy and identify a threshold of 38%. We use this threshold for all the patients in our clinical study.

Obstructive apnea estimation algorithm. As described earlier we detect obstructive apnea using sudden spikes in the chest motion signal. We identify the spikes from the peaks in the signal when the amplitude in these peaks increases by 50%. Further, if multiple spikes occur within a second we consider them to be part of a single obstructive apnea event. We note that some of the patients with periodic limb movement disorder periodically move their limbs. To avoid confusing them with obstructive apnea, we ignore periodic spikes in the chest motion signal.

3.2.2 Estimating the total sleep time

In a polysomnography test, the EEG sensors are used to measure the brain activity to determine whether the subject is asleep or awake; this is then used to measure the total sleep time of a subject. Recent work [30] has shown that one can use body movements to compute the total sleep time and the resulting accuracies are acceptable for the purposes of estimating the apnea-hypopnea index. Thus, ApneaApp measures the total sleep time by identifying the non-breathing body movements and subtracting their duration from the total experiment time.

To identify these movements, we leverage that when the subject's body moves, the sonar reflections experience large variations that do not have the periodicity of the breathing motion. By identifying these aperiodic signals that are shown in Fig. 7, ApneaApp detects body movements. Specifically, we run our peak detection algorithm after disabling the minimum distance heuristic. We then check if the resulting peaks are aperiodic. To do this, we perform a second order derivative (difference of the difference) over the peak locations. For periodic breathing motion, this derivative is close to zero but is higher for aperiodic signals. We use a threshold of five to detect such aperiodic signals and compute the corresponding body motion durations.

Now, we compute the total sleep time by subtracting the body motion durations from the total experiment time. We note that if a patient moves twice within a short period, it is likely that she was awake for the whole duration between the movements. Thus, for the purposes of computing the sleep time, we combine any body movements that occur within a ten-minute period into a single awake duration.

4. IMPLEMENTATION

We implement ApneaApp as a third party Android app that does not require rooting the smartphone. ApneaApp plays frequency-

	Samsung S4	S5	Galaxy Nexus	HTC One
18-20 kHz Mic	Yes	Yes	Yes	Yes
18-20 kHz Speaker	Yes	Yes	Yes	Yes
Separated Speaker/Mic	Yes	Yes	No	Yes
Apnea Detection	Yes	Yes	No	Yes
Breathing Frequency	Yes	Yes	Yes	Yes

Table 1—Smartphone requirements for ApneaApp. The table shows that accurate apnea detection requires a phone that has a speaker-microphone pair that is not co-located. This is because a co-located speaker-microphone pair results in unpredictable variations in the amplitude of the microphone signal. However, monitoring breathing frequency does not require stable amplitude and hence can work even with co-located speakers and microphones.

modulated samples on the phone speaker and continuously record the raw samples from the microphone.

To understand the smartphone capabilities necessary to accurately assess sleep apnea, we experiment with four different Android phones: Samsung Galaxy S4, Samsung Galaxy S5, HTC One, and Galaxy Nexus. To operate ApneaApp the phone requires the ability to transmit and receive audio signals in the 18-20 kHz range from the speaker and microphone respectively. Table. 1 shows that all the tested smartphones satisfy this requirement.

Another critical requirement is a stable amplitude at the microphone. This is important since our algorithm leverages amplitude changes to detect sleep apnea events. Our experiments reveal that, across all the tested smartphones, when the co-located microphone-speaker pair concurrently transmit and receive signals, the amplitude at the microphone experiences unpredictable variations. This happens because of a feedback loop between the co-located microphone and speaker and can result in sleep apnea misclassifications.

To address this issue, we leverage that many smartphones today come with an additional microphone that is not co-located with the speaker. Our results show that using this microphone eliminates any unintended interaction with the phone speaker. This means that smartphones that do not have this additional microphone, such as the Galaxy Nexus, cannot be used to accurately detect sleep apnea events with ApneaApp. Note, however, that the measurement of breathing frequency is not affected by this feedback loop and hence ApneaApp can still be used to track the breathing frequency on Galaxy Nexus phones.

5. CLINICAL STUDY

We conducted a clinical sleep study with 37 patients (17 female and 20 male) between ages of 23-93 (mean: 50) for a total of 296 hours and compare the results from ApneaApp with the PSG study reports. The study was conducted at the sleep laboratory in the UW Medicine Sleep Center at Harborview.

Participants. The patients who participated in our study were examined by a sleep physician for common sleep disorders like excessive snoring or daytime fatigue and prescribed to undergo the in-laboratory polysomnography (PSG) study. The sleep lab is equipped to conduct PSG studies for a maximum of eleven patients per night; we randomly choose up to five subjects per night for our study. We do not screen patients based on their gender, race or national origin, but only consider adults. The patients who participated in our study were not provided any monetary benefits. We note two key points: First, our participant group included both patients undergoing the regular PSG study to diagnose sleep apnea as well as those undergoing the CPAP treatment [27] after being diagnosed positive for sleep apnea. Second, out of the 38 patients we approached, 37 consented to participate in our sleep study. Further,

there was significant interest amongst them in using a smartphone instead of the existing PSG procedure to diagnose their sleep illness.

Protocols. In clinical PSG studies, the sleep lab technician assigns each patient to a separate room furnished with an adjustable king size bed. Each room is also fitted with a camera to get a visual log of the patient. Polysomnography was performed as follows according to American Academy of Sleep Medicine criteria: Lectro-encephalographic electrodes were positioned at two frontal (F7, F8), two central (C3, C4), and two occipital (O1, O2) locations (International 10-20 system of measurement) and were referenced to the contralateral mastoids. Chin electromyogram and right and left electroculogram electrodes were also applied. Airflow was measured using a nasal pressure cannula placed in the nose and a thermistor placed in the nose and over the mouth, allowing differentiation between nasal and oral breathing. Chest and abdominal respiratory effort were assessed by piezo respiratory-effort bands placed around the chest and abdomen, and snoring with a small microphone sensor placed on the throat just lateral to the trachea (Pro-Tech Services, Inc., Mukilteo, WA). Oxygen saturation was measured from the index finger via pulse oximetry (Nellcor, Pleasanton, CA). Bilateral electromyogram electrodes were placed on the anterior tibialis muscle to monitor leg movements. After fitting the sensors, the technician performs bio-calibrations by asking the patient to do specific activities like leg motion, snoring and breathing exercises and ensure that the sensors are properly attached.

The procedures for our ApneaApp study occur concurrently with existing PSG protocols. Specifically, as the sleep technician attaches sensors to the patient, the first author places a Samsung Galaxy S4 smartphone running ApneaApp at the other end of the bed. The position and the orientation of the phone vary across patients depending on their sleep habits. For example, some patients prefer to surround themselves with pillows in which case the phone is placed behind the pillows at a farther distance. The distance of the phone from the patient varies between 0.3 and 0.7 meters. We leave the room and return in the morning to collect the smartphones from the sleep lab. The phone is plugged in throughout the study.[5] The above procedures were reviewed and approved by the UW Human Subjects Division.

Data Processing. The data from the clinical PSG study is sent to a third-party entity that scores the sensor data and provides the number of central and obstructive apneas, and hypopneas. Currently, a trained sleep technician manually performs the above scoring process resulting in a minimum delay of three days to obtain the scored data. Specifically, obstructive apneas, hypopneas, and central apneas were scored by the technician as follows: Obstructive apneas were defined by at least a 90% reduction in the pressure-flow signal with corresponding respiratory effort; central apneas were defined by at least a 90% reduction in pressure-flow signal without respiratory effort; and hypopneas were defined as a greater than 30% reduction in amplitude in pressure-flow signal lasting at least 10 seconds with an associated 4% oxygen desaturation.

To process the sonar data collected using ApneaApp, we run the algorithm described in §3.2. The scored ground truth data per pa-

---

(a) No.of. Central Apnea Events

(b) No.of. Hypopnea Events

(c) No.of. Obstructive Apnea Events
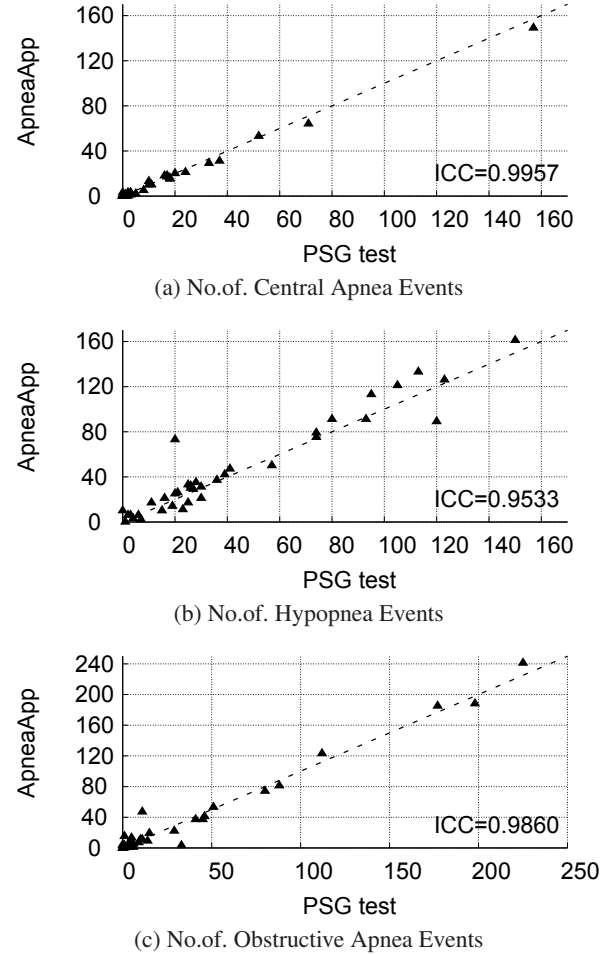
Figure 8—Scatter-plots of the Number of Apnea Events. The x-axis and the y-axis show the number of apnea events estimating by PSG and ApneaApp respectively. The plots show that there is a high correlation between the number of central, obstructive and hypopnea events identified by PSG and ApneaApp.

tient from the PSG study is more than 1 GB. Thus, we compare it with ApneaApp's signals using Matlab.

5.1 Sleep Apnea Detection Accuracy

The key metric used for sleep apnea diagnosis is the Apnea-Hypopnea Index (AHI), which represents the average rate at which apnea and hypopnea events occur during the sleep duration. To compute this, as described in §2, we need to estimate the number of central, obstructive and hypopnea events as well as the total sleep time. In this section, we evaluate ApneaApp's effectiveness in estimating these parameters.

Accuracy of apnea events. First we examine ApneaApp's effectiveness in computing the number of central apnea, hypopnea, and obstructive apnea events. We compare the number of apnea events computed by our algorithm described in §3.2 with the baseline numbers provided by the PSG study. Figs. 8 (a)-(c) show the scatter plots for central, obstructive and hypopnea events. The 45-degree line denotes the ideal scenario when our estimates match the baseline.

The plots show that the number of apnea events detected by ApneaApp is highly correlated with the PSG test. Specifically, the in-

False Positives	False Negatives
0.0023% ($\frac{59}{26070}$)	0.0336% ($\frac{43}{1281}$)

Table 2—Detecting Body Motion. ApneaApp detects body motion using abrupt changes to the periodicity of breathing movements. We use the body motion sensors from the PSG study as our ground truth. The table shows that we can accurately detect body movements without any sensor instrumentation of the human body.

terclass correlation coefficient (ICC)[6] between ApneaApp and PSG is 0.9957, 0.9533 and 0.9860 for central apnea, hypopnea and obstructive apnea respectively. We further note the following.

- Amongst the apneas, the number of central apnea events computed by ApneaApp has the highest correlation (ICC=0.9957). This is because a central apnea event is detected as the absence of chest motion; this can be done accurately using the algorithm in §3.2.
- The number of hypopneas has a lower correlation coefficient (ICC=0.9533). The key reason for this is that hypopneas are detected when the amplitude of the breathing motion reduces by more than a threshold. Thus, small errors near this threshold can result in misclassifications. We note that in the PSG study, the technicians mark the hypopnea events manually by ball parking the 30% threshold; this introduces a number of misclassified hypopnea events in the ground truth data itself. We also note that the error due to the lack of oxygen desaturation information in the case of ApneaApp was negligible.
- ApneaApp can accurately compute the number of obstructive apneas by monitoring the chest movements alone (ICC =0.9860), without a nasal pressure transducer that can directly monitor air flow. Analyzing the misclassifications further reveals that ApneaApp misses obstructive apnea events where the airflow is obstructed for a very small duration (10-15s). This is expected because such events might result in less noticeable changes in the breathing effort/movements but would be noticeable at the nasal pressure transducer signal used in the PSG test.

Body Movements and sleep time estimation. To compute AHI, we need to measure the total sleep time. In the clinical PSG study, the staging step in the scoring process calculates the total sleep time using the brain activity (EEG) information. Since ApneaApp does not have the brain activity data, it measures the sleep time indirectly using body movements: the absence of body movements is used as an indicator that the patient is asleep. We evaluate ApneaApp's accuracy first in detecting non-breathing body movements and then in estimating the sleep time.

(i) Body Motion Detection Accuracy. ApneaApp detects non-breathing body movements using abrupt changes in the periodicity of the breathing movements (see §3.2.2). To evaluate this, we use the body motion sensor data from the PSG study as the ground truth. Every 30 second epoch in the PSG data is annotated with the movement information; we consider each of these epochs as an event. We count a false positive when a body movement is detected by ApneaApp in the absence of one and a false negative when Ap-

---

[6]ICC is a standard statistic used in the sleep literature for quantifying the effectiveness of two procedures (in our case, PSG and ApneaApp) to produce correlated observations across a range of values [30]. ICC is also a more appropriate metric for us compared to recall accuracy: missing one of four apnea events results in a low recall accuracy of 75%. This however does not capture the system-level performance since for a eight-hour duration, the resulting AHI error is only 0.125.
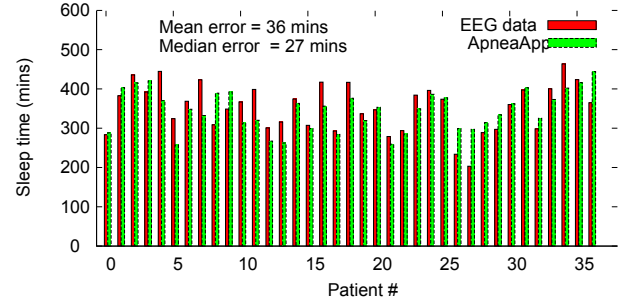
Figure 9—Sleep time Estimation Error. We compare the sleep time estimated using body movements (ApneaApp) and brain activity (EEG data). The mean and median error is 36 and 27 mins. ApneaApp tends to overestimate the sleep time for patients who wake up and lie on the bed without any movements while trying to fall asleep. ApneaApp's sleep time accuracies are however sufficient for the purposes of diagnosing sleep apnea. We also note that in practice patients could provide direct feedback about these unusually long awake periods to improve our estimation further.

neApp does not detect a body movement that is present in the PSG data.

Table. 2 shows the number of false positive and false negatives across the 37 patients. The results show that ApneaApp can accurately detect 1238 of the 1281 body movement epochs. Further, ApneaApp misclassifies only 59 epochs as body movements across all the 37 patients. This is expected because body movements create significant changes to the reflected signal and hence can be easily distinguished from breathing movements. Thus we conclude that ApneaApp can accurately detect body movements without the need for placing sensors on the human body.

(ii) Sleep Time Estimation Accuracy. ApneaApp computes sleep time by subtracting the total duration of body movements from the total experiment time. Next, we evaluate how well this correlates with the sleep time computed in the PSG study using the brain activity (EEG sensors). Fig. 9 plots the total sleep time computed from both ApneaApp and the EEG sensor data across the 37 subjects. The plot shows that the mean and median sleep time error is 36 and 27 mins respectively. Further, ApneaApp overestimates the sleep time by an hour for six patients. This happens because these patients woke up in the middle of the night and lie on the bed without frequent movements, while trying to fall asleep. The brain activity during this period is high but there is little to no frequent motion. This is a fundamental limitation of estimating the total sleep time without access to the brain activity data. However, as we will see next, ApneaApp's sleep time accuracies are acceptable for diagnosing sleep apnea.

Accuracy of AHI estimation. Finally, we evaluate how well ApneaApp can compute the apnea-hypopnea index and diagnose sleep apnea. The results are as follows:

- Fig. 10 plots a scatter plot of the AHI values computed by ApneaApp versus PSG for the 37 patients. The figure shows that the AHI values are highly correlated between ApneaApp and PSG with an interclass correlation coefficient of 0.9816. Further, the average error in computing AHI is 1.9 events/hr. Larger errors tend to happen at higher AHI values where the diagnosis is fortunately more tolerant to large errors, i.e., the diagnosis is the same whether the AHI is 40 or 50.
- Fig. 11 shows the confusion matrix of the sleep apnea levels diagnosed by ApneaApp with thresholds of 5, 15, and 30. The ta-
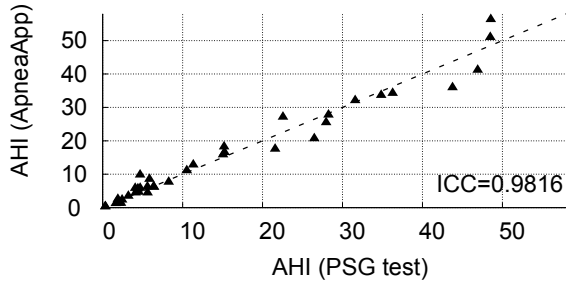
Figure 10—Scatter-plot of the AHI Values. The mean error in the AHI values computed between PSG and ApneaApp is 1.9. Larger errors typically happen at higher AHIs where the diagnosis is fortunately more tolerant to large error, i.e., the diagnosis is the same whether the AHI is 40 or 50.
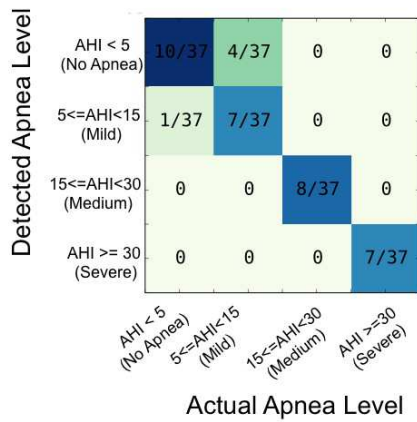

Figure 11—Confusion Matrix of the Sleep Apnea Levels. We accurately classifies 32 out of 37 patients between four sleep apnea levels. The five misclassifications occur between no-apnea and mild-apnea; four of them happen right at the boundary between the two levels with an error less than 1 event/hr. These boundary cases are handled separately by physicians depending on the patient preferences, symptoms, and insurance, effectively reducing the number of misclassifications to one.

ble shows that ApneaApp diagnoses the sleep apnea level accurately for 32 out of 37 patients. The five misclassifications are between the no-apnea and mild-apnea diagnosis. Specifically, four patients who have mild-apnea are diagnosed by ApneaApp as having no apnea while one patient with no-apnea is diagnosed as having mild sleep apnea.

- For four of the five misclassified patients, the AHI values are in the 4.5-5.5 range centered at the threshold between no-apnea and mild-apnea. These boundary cases are handled separately by physicians depending on the patient preferences, symptoms, and insurance, effectively reducing the number of misclassifications to one.

6. MICROBENCHMARKS

We run experiments with five participants (ages 24–25) in a home bedroom environment to evaluate the effects of various parameters such as orientation, position and distance. The bedroom is next to a major street with significant foot and vehicular traffic. Since we cannot imitate the sleep apnea events where the chest motion experiences amplitude variations, in these set of experiments we use the
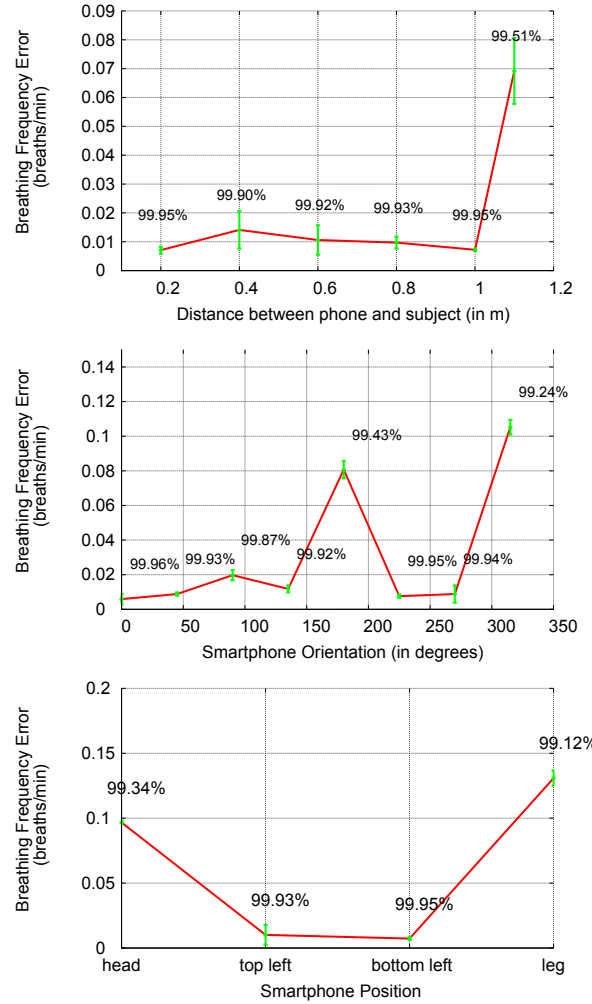
Figure 12—Effects of phone distance, orientation and position. The plots show the results for a subject sleeping in the supine position.

coarse-grained breathing frequency as a proxy to understand the effects of the various parameters. To obtain the ground truth data for the breathing frequency, we use a Vernier respiratory belt worn at the abdomen level.

6.1 Effect of Phone Distance, Orientation and Position

We first evaluate the effect of the phone distance, orientation and position on the breathing frequency accuracies.

Effect of the phone's distance. We place the phone on the left side of the subject sleeping in the supine position. In each trial, we monitor the breathing frequency over a three-minute duration after which the subject is asked to get out of the bed. For each distance value, we perform five trials for a total of ten minutes per location. We measure the accuracy of the breathing frequency estimated by ApneaApp versus the Vernier respiratory belt and plot the results in Fig. 12. The figure shows that:

- The breathing frequency accuracy is as high as 99.90% at distances up to one meter. We note that these accuracies are much higher than those observed for sleep apnea. This is because, as

Figure 13—The various phone positions used for the results in Fig. 12. We experiment with four different positions along a semicircle centered at the subject.

Figure 14—Effect of sleeping position. The accuracy is lower when the patient is lying with her face down (prone). In this position, both the signals from the Vernier belt and ApneaApp experience larger variations. We however note that in our clinical study we track the chest movements throughout the sleep duration where the patient's sleeping position was not controlled.

Figure 15—Effect of blankets. We use four blankets with thicknesses varying from 2-5 cm. The plot shows that the accuracies are high even when blankets separate the subject from the phone.

described in §4, the breathing frequency on the phone is relatively stable compared to the amplitude variations that are necessary for detecting sleep apnea events.

- As the distance increases beyond a meter the accuracies decrease. This is because the strength of the reflections due to breathing reduces with distance, making the breathing signal noisy. The one-meter range is, however large enough to enable contactless breath monitoring that is non-intrusive, as demonstrated in our clinical study. It also limits the negative effects of environmental changes farther away than a meter on ApneaApp's accuracies.
- The accuracies are unaffected by audible noise in the environment from the vehicular and foot traffic on the street. Introducing human conversations in the vicinity of the experiments also does not affect these accuracies. This is because, we use a high-pass filter to filter out audible signals below 18 kHz.
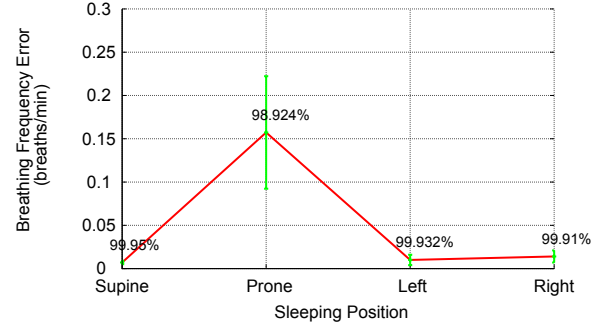
Effect of the phone's orientation. We place the phone 20 cm away and to the left of the subject. We then rotate the phone and compute the accuracies for each phone orientation. As before, for each trial, we perform five experiments for a total of 10 minutes per phone orientation. Fig. 12 plots the results as a function of the phone's orientation. We observe that the accuracies remain high, demonstrating that during ApneaApp's operation we do not need to fix the phone orientation.

Effect of the phone's position. Next, we experiment with the phone at different positions around the subject. Specifically we place the phone is four different positions — near the head, near the legs, and two positions to the left — along a semicircle of radius 40 cm centered at the subject as shown in Fig. 13. Fig. 12 shows that the accuracies are high when the phone is in the left positions and slightly lower when placed near the head and the feet. This is because in the latter positions, the head and the leg effectively block the chest/abdomen motion. We however note that the maximum error is less than 0.13 breaths per minute across all the phone positions.
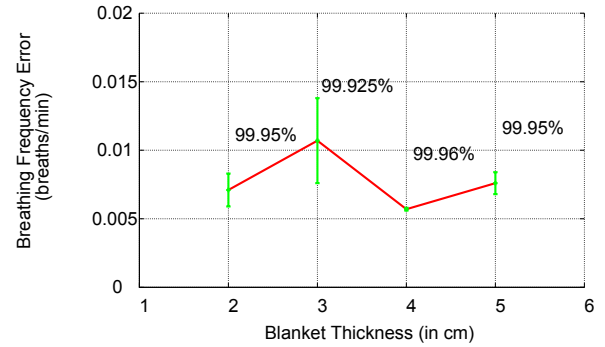
6.2 Effect of Sleeping Position and Blankets

Next, we evaluate the accuracies for different sleeping positions and in the presence of blankets.

Effect of the subject's sleeping position. We consider four different sleeping positions: supine (on the back), prone (on the abdomen),

lying on the left, and the right. We place the smartphone at a distance of 20 cm to the left of the subjects and measure the breathing rate accuracies. As before, for each sleep position we monitor the breathing rate over chunks of two minutes for a total of ten minutes. Fig. 14 shows that the average residual error is below 0.16 breaths per minute across all the sleeping positions. We note that the accuracy is lower when the patient is lying with her face down (prone). In this position, we noticed that both the signals from the Vernier belt and ApneaApp experience a larger variation. We however note that our clinical study tracks the chest movements throughout the sleep duration where the patient's sleeping position was not controlled.

Effect of Blankets. We measure the breathing frequency accuracy for various blanket thicknesses. The subjects are asked to sleep in the supine position and the phone is placed left of the subject at a distance of 40 cm. We use four blankets with thicknesses varying from 2-5 cm. Fig. 15 shows that the accuracies are not noticeably degraded by the use of blankets. This demonstrates that ApneaApp is well suited for the sleep environment, which is further validated by our clinical study where all the patients used blankets.

6.3 Breathing Signals from Multiple Subjects

As discussed in §3.1, the sonar reflections from multiple subject arrive at different times at the microphone. Thus, ApneaApp can simultaneously track breathing movements from more than one sub-
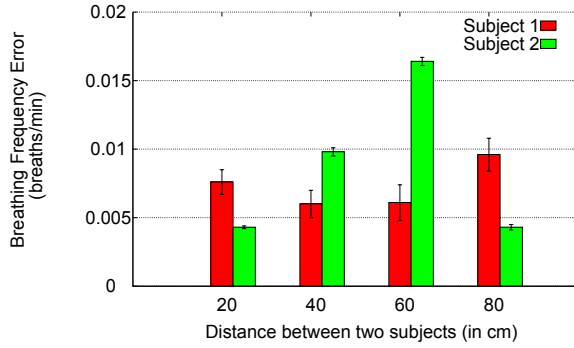
Figure 16—Extracting breathing from two subjects. The plot shows that ApneaApp can simultaneously track and monitor the breathing signals from two subjects on the bed.

ject. To evaluate this, we monitor the breathing rate from two subjects sleeping in the supine position on the bed. Both the subjects where fitted with the Vernier respiratory strap to record the ground truth. The phone was placed at one side of the bed at a distance of 20 cm from the first subject. We increase the distance between the two subjects at intervals of 20 cm starting at the minimum distance of 20 cm where the two subjects could sleep without overlapping. As before, we repeat the experiments over five trials for a total of 15 minutes at each distance value.

Fig. 16 plots the accuracies for the two subjects as a function of the distance between them. The plots show that ApneaApp can accurately distinguish between breathing signals of subjects even when they are as close as 20 cm from each other. This is expected, because as described in §3.1, our FMCW transmissions can be used to resolve reflections with a greater resolution than 20 cm. The key point, however, is that we can identify and track the breathing signals from more than one subject — a capability that is important for use in the home environment.

6.4 Evaluating ApneaApp's Audibility

ApneaApp transmits FMCW audio signals in the 18-20 kHz range. To evaluate the audibility of these signals, we ran ApneaApp with 144 subjects: 37 of whom were sleep apnea patients who took part in our clinical study, 50 additional sleep apnea patients not part of our clinical study, and 57 healthy undergraduate subjects who are seniors in the UW CSE department.

None of the 87 sleep patients reported any audible sounds from the phone. Further, the 37 patients who took part in our clinical study had no audibility issues for the eight-hour sleep duration. One of the undergraduates could hear faint sounds when the phone was placed next to his ear and a second undergraduate reported hearing a sound similar to TV static at distances of up to one meter. However 55 of the 57 undergraduate students reported no audible sounds from the phone even when placed right next to the ear. This shows that ApneaApp is inaudible for most adults, which is the target population for ApneaApp.

7. RELATED WORK

Our work is related to prior art in three domains.

(a) Mobile Health and Wellness Systems: There has been recent interest in the mobile health and wellness community to better understand sleep quality. Wearable sleep sensors such as FitBit [2], WakeMate [14], and Jawbone [5] capture the movements from an accelerometer to sense basic sleep patterns such as hours slept and

the number of wakings. Lullaby [26] leverages sensors deployed in the bedroom to understand the effect of environmental factors (e.g., ambient noise and lighting conditions) on sleep quality. More recent work including Toss-N-Turn and iSleep leverage smartphone sensors as sleep quality detectors [29, 19, 24]. Specifically, they use accelerometers to detect coarse non-breathing body movements [29, 24] and microphones for cough [24] and snoring [24, 11]. ApneaApp builds on this foundational work but is complementary to it. We introduce a novel contactless system for smartphones that can track fine-grained chest and abdomen movements due to breathing. We then demonstrate the ability to detect sleep apnea events on smartphones.

Our work is also related to prior work on monitoring physiological signals using wearable sensors [12] including mouthpieces and chest bands [13, 7]. Designs such as BodyBeat [34] use custombuilt piezoelectric microphones on the body surface to monitor body sounds such as respiration, food intake and laughter. More recent work [36] leverages audible breathing sounds from an ear bud to infer the breathing frequency. The Android app, iBreath [23], combines the audible breathing sounds from a Bluetooth headset with the accelerometer on the phone to infer the breathing frequency while running. These systems are very promising for their application domains. For diagnosing sleep apnea, however, we are unaware of clinical-validated research that demonstrates the ability to compute AHI values and detect sleep apnea events using only audible breathing sounds/frequency. In contrast, we propose a novel contactless solution that can monitor the chest movements by actively transmitting FMCW audio signals. We then demonstrate the ability to detect various sleep apnea conditions including obstructive, central, and hypopnea as well as compute the AHI values.

(b) RF-based Breathing Rate Monitors: Prior RF-based approaches [44, 37, 22] leverage changes in radio signals to estimate the breathing frequency. This includes using a network of 2.4 GHz wireless sensors in the bedroom [31, 32] as well as expensive ultrawideband RF radar systems with 500 MHz to 2 GHz of bandwidth [35, 45]. While promising, none of these approaches have been demonstrated to work on off-the-shelf smartphones. We also believe that the sonar approach introduced in this paper is more attractive for two main reasons. First, a microphone that samples at 48 kHz gives a range resolution of 0.7 cm; RF based approaches need Gigahertz of bandwidth to get a similar resolution. Second, the one-meter operational range of our sonar system reduces the effect of environmental movements making it more robust in real world sleep scenarios.

(c) Sleep Apnea Research and Systems: Polysomnography [18] is the standard clinical test for diagnosing sleep apnea in the medical setting. This however is an expensive test that requires attaching various sensors to the patient for the sleep duration. While portable recording systems [25, 20] are being developed for use in home settings, they require the attachment of various sensors and most still require a trained technician to setup the system [30].

Mobile phone app developers have also designed questionnaires (e.g., Home Sleep Apnea A-Z [3], Sleep Access [8]) to screen patients for the PSG test. Sleep Apnea Monitor app [10] help users diagnosed with sleep apnea to sleep on their side instead of their back by buzzing when the user moves to their back. It does so by attaching the phone to the person's leg. Prior research has also explored the feasibility of leveraging snoring and movement information [15, 17] to perform screening for PSG. While correlation existing between sleep apnea and snoring, such an approach does not monitor breathing movements and hence cannot identify sleep

apnea events [42, 30]. To address this issue, commercial systems such as Sonomat [30] use embedded sensors in a mattress pad to track the breathing movements and sounds [1, 30]. Our work instead focuses on detecting sleep apnea events by tracking breathing movements on smartphones, without the need for any other sensors.

Finally, medical research has proposed the use of ultrasound imaging, instead of MRI, to measure the tongue volume and width of upper airway [16, 38]. This information has been demonstrated to be helpful in diagnosing obstructive apnea [28]. In contrast to this work, we transform a smartphone into an active sonar system and develop algorithms to extract hypopnea, obstructive apnea, and central apnea information from the sonar signal changes.

8. LIMITATIONS AND FUTURE DIRECTIONS

We discuss the limitations of our system as well as various opportunities to improve it.

Increasing the AHI estimation accuracies. We design ApneaApp to operate in a home setting without the need for instrumenting the users with sensors. We note however that a clinical PSG test uses a number of sensors that provide physiological information that is not available in our current design. For example, EEG sensors provide the brain activity data that helps accurately find different sleep stages (REM, non-REM, and awake). One could imagine analyzing fine-grained changes in the sonar reflections to find correlations with the sleep stages. Specifically, we plan to explore correlations between physiological parameters such as breathing rate and motion with the various sleep stages.

Similarly, since the PSG test uses cameras to produce a visual log of the patient, the report also summarizes the patient's sleeping position and the corresponding number of sleep apnea events. We plan to explore the feasibility of using accelerometers to predict these sleeping positions. Another direction is to augment the chest movements from our system with audible audio signals including snoring and evaluate the resulting accuracy improvements. Supplementing our system with these capabilities can improve our accuracies and also provide additional diagnostic information.

Detecting other respiratory-related events. In addition to diagnosing sleep apnea, a PSG test is also used to detect other events such as the respiratory-effort related-arousals (RERAs). These are arousals in the sleep that do not meet the definition of apnea and hypopnea events. A direction worth exploring is to leverage our sonar design with other smartphone sensors to detect these non-sleep apnea conditions.

Extracting other physiological information. In this paper, we focus on extracting the minute chest and abdomen movements due to breathing. We note however that in principle our sonar design can also be used to extract the heart rate without requiring contact with the human body. Prior RF-based radar systems have demonstrated the feasibility of extracting heart rate from radio signals with Gigahertz of bandwidth. While heart rate is not necessary for sleep apnea diagnosis, we plan to use narrowband sonar transmissions on a smartphone and enable contactless heart rate detection, with the goal of better understanding the sleep architecture [40].

9. CONCLUSION

We present a contactless solution that detects sleep apnea events on smartphones. To achieve this, we introduce a novel technique that can track chest and abdomen movements due to breathing on smartphones using FMCW sonar transmissions. We design algorithms to detect central apnea, obstructive apnea, and hypopnea as

well as estimate the total sleep time by analyzing the FMCW reflections from the human body. We implement our design on off-the-shelf smartphones and demonstrate the ability to concurrently track breathing movements from multiple subjects. Finally, we present results from a clinical study with 37 patients that demonstrate the feasibility of using our system to accurately compute the number of central, obstructive, and hypopnea events as well as the apnea-hypopnea index.

10. REFERENCES

[1] Bam labs, sleep iq only from sleep number. http://bamlabs.com.

[2] Fitbit. www.fitbit.com.

[3] Home sleep apnea a-z. https://play.google.com/store/apps/details?id=com.app_aviisha123.layout&hl=en.

[4] Iphone app: Respiratory rate. https://itunes.apple.com/us/app/respiratory-rate/id538726900?mt=8.

[5] Jawbone. https://jawbone.com/.

[6] Morbidity and mortality weekly report,centers for disease control and prevention. In Weekly, Vol.60, No. 8, 2011.

[7] Nose breathe mouthpiece. http://www.nosebreathe.com/.

[8] Sleep access. http://sleepaccess.com/.

[9] Sleep apnea: Epidemiology. http://www.sleepdex.org/epidemiologyapnea.htm.

[10] Sleep apnea monitor. https://itunes.apple.com/gb/app/sleep-apnea-monitor/id464587229?

[11] Sleep as android: Snoring detection. https://sites.google.com/site/sleepasandroid/.

[12] Snuza - portable baby monitors. http://www.snuza.com/.

[13] Vernier respiration monitor belt. http://www.vernier.com/products/sensors/rmb/.

[14] Wakemate. https://play.google.com/store/apps/details?id=com.wakemate.android.

[15] U. Abeyratne, S. Silva, C. Hukins, and B. Duce. Obstructive sleep apnea screening by integrating snore feature classes. In Physiological Measurement, 2013.

[16] M. Al-Abed, P. Antich, D. Watenpaugh, and K. Behbehani. Detection of airway occlusion in simulated obstructive sleep apnea/hypopnea using ultrasound: An in vitro study. In Engineering in Medicine and Biology Society (EMBC), 2013.

[17] S. Alqassim, M. Ganesh, S. Khoja, M. Zaidi, F. Aloul, and A. Sagahyroon. Sleep apnea monitoring using mobile phones. In e-Health Networking, Applications and Services (Healthcom), 2012.

[18] I. C, A.-I. S, C. A. Jr., and Q. SF. The aasm manual for the scoring of sleep and associated events. In 1st ed. Westchester IL: American Academy of Sleep Medicine, 2007.

[19] Z. Chen, M. Lin, F. Chen, N. D. Lane, G. Cardone, R. Wang, T. Li, Y. Chen, T. Choudhury, and A. T. Campbell. Unobtrusive sleep monitoring using smartphones. In PervasiveHealth 2013.

[20] W. J. Deutsch PA, Simmons MS. Cost-effectiveness of split-night polysomnography and home studies in the evaluation of obstructive sleep apnea syndrome. In J Clin Sleep Med 2006;2:145-53.

[21] W. W. Flemons, N. J. Douglas, S. T. Kuna, D. O. Rodenstein, and J. Wheatley. Access to diagnosis and treatment of patients with suspected sleep apnea. In American Journal of Respiratory and Critical Care Medicine, Vol. 169, No. 6 (2004).

[22] N. Fox, C. Heneghan, M. Gonzalez, R. Shouldice, and P. de Chazal. An evaluation of a non-contact biomotion sensor with actimetry. In Engineering in Medicine and Biology Society, 2007.

[23] T. Hao, G. Xing, and G. Zhou. ibreath: A convenient mobile app that tracks breathing during running.

[24] T. Hao, G. Xing, and G. Zhou. isleep: Unobtrusive sleep quality monitoring using smartphones. In SenSys '13.

[25] M. JF, C. J, and e. a. Pereira R. Effectiveness of home respiratory polygraphy for the diagnosis of sleep apnoea and hypopnoea syndrome. In Thorax 2011;66:567-73.

[26] M. Kay, E. K. Choe, J. Shepherd, B. Greenstein, N. Watson, S. Consolvo, and J. A. Kientz. Lullaby: A capture & access system for understanding the sleep environment. In UbiComp '12.

[27] C. A. Kushida, A. Chediak, R. B. Berry, L. K. Brown, D. Gozal, C. Iber, S. Parthasarathy, S. F. Quan, and J. A. Rowley. Clinical guidelines for the manual titration of positive airway pressure in patients with obstructive sleep apnea. Journal of Clinical Sleep Medicine.

[28] Y. Lahav, E. Rosenzweig, Z. Heyman, J. Doljansky, A. Green, and Y. Dagan. Tongue base ultrasound: a diagnostic tool for predicting obstructive sleep apnea. In Ann Otol Rhinol Laryngol, 2009.

[29] J.-K. Min, A. Doryab, J. Wiese, S. Amini, J. Zimmerman, and J. I. Hong. Toss 'n' turn: Smartphone as sleep and sleep quality detector. In CHI '14.

[30] M. Norman, S. Middleton, O. Erskin, P. Middleton, J. Wheatley, and C. Sullivan. Validation of the sonomat: a contactless monitoring system used for the diagnosis of sleep disordered breathing. In Sleep, Sep 2014.

[31] N. Patwari, L. Brewer, Q. Tate, O. Kaltiokallio, and M. Bocca. Breathfinding: A wireless network that monitors and locates breathing in a home. Selected Topics in Signal Processing, IEEE Journal of, 2014.

[32] N. Patwari, J. Wilson, S. Ananthanarayanan, S. Kasera, and D. Westenskow. Monitoring breathing via signal strength in wireless networks. Mobile Computing, IEEE Transactions on, 2014.

[33] Rafael Golpe and Antonio JimÃl'nez and Rosario Carpizo. Home sleep studies in the assessment of sleep apnea/ hypopnea syndrome. In CHEST, 2002.

[34] T. Rahman, A. T. Adams, M. Zhang, E. Cherry, B. Zhou, H. Peng, and T. Choudhury. Bodybeat: A mobile system for sensing non-speech body sounds. In MobiSys '14.

[35] T. Ralston, G. Charvat, and J. Peabody. Real-time through-wall imaging using an ultrawideband multiple-input multiple-output (mimo) phased array radar system. In Phased Array Systems and Technology (ARRAY), 2010.

[36] Ren, Yanzh and Wang, Chen and Chen, Yingying and Yang, Jie. Poster: Hearing your breathing: fine-grained sleep monitoring using smartphones. In MOBICOM Posters, 2014.

[37] R. Shouldice, C. Heneghan, G. Petres, A. Zaffaroni, P. Boyle, W. McNicholas, and P. de Chazal. Real time breathing rate estimation from a non contact biosensor. In Engineering in Medicine and Biology Society (EMBC), 2010.

[38] C.-C. Shu, P. Lee, J.-W. Lin, C.-T. Huang, Y.-C. Chang, C.-J. Yu, and H.-C. Wang. The use of sub-mental ultrasonography for identifying patients with severe obstructive sleep apnea. PLoS ONE, 2013.

[39] K. Strohl, S. Merritt, J. Blatt, A. Pack, F. Council, S. Rogus, K. Georges, T. Roth, J. Kiley, j Stutts, R. Kurrus, P. Waller, A. McCartt, and Willis. Ncsr/nhtsa epert panel on driver fatigue and sleepiness. In Drowsy Driving and Automobile Crashes.

[40] K. Tanida, M. Shibata, and M. Heitkemper. Sleep stage assessment using power spectral indices of heart rate variability with a simple algorithm: limitations clarified from preliminary study. Biol Res Nurs, 2013.

[41] Terry Young and Mari Palta and Jerome Dempsey and James Skatrud and Steven Weber and Safwan Badr. The occurrence of sleep-disordered breathing among middle-aged adults. In The New England Journal of Medicine, 1993.

[42] A. Visvanathan, A. Gibb, and R. Brady. Increasing clinical presence of mobile communication technology: avoiding the pitfalls. In Telemed J E Health, 2011.

[43] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman. Eulerian video magnification for revealing subtle changes in the world. ACM Trans. Graph., 2012.

[44] A. Zaffaroni, P. de Chazal, C. Heneghan, P. Boyle, P. Mppm, and W. McNicholas. Sleepminder: an innovative contact-free device for the estimation of the apnoea-hypopnoea index. In Proc. IEEE Engineering Medical biology Society, 2009.

[45] D. Zito, D. Pepe, M. Mincica, and F. Zito. A 90nm cmos soc uwb pulse radar for respiratory rate monitoring. In Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011.