Applications of the Analysis of Respiratory Kinematics (ARK) System to Respiratory Distress Outcome Prediction

---

A Dissertation Presented to
the Faculty of the School of Engineering and Applied Science
University of Virginia

In Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy
Computer Engineering

by

William Ashe

May 2022

Approval Sheet

This dissertation is submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Computer Engineering

William Ashe

This dissertation has been read and approved by the Examining Committee:

Ronald D. Williams, Advisor

Joanne Bechta Dugan

Aidong Zhang

J. Randall Moorman

Shrirang Gadrey

Accepted for the School of Engineering and Applied Science:

Jennifer L. West, Dean
School of Engineering and Applied Science

May 2022

Abstract

Clinical respiratory monitoring has been historically limited by the available technology that can reliably produce useful metrics without jeopardizing patient comfort. Pulmonary condition can be accurately assessed by physicians who have the training to identify relevant warning signals from visual examinations and available metrics such as average respiratory rate and oxygen saturation. Traditional methods of laboratory pulmonary function testing are powerful and diverse, but lab procedures would generally be infeasible for widespread application due to the cumbersome setup and equipment and the time required to take measurements. To combat this translational research gap, new methods using lightweight technologies like wearable sensors have proliferated, demonstrating large potential for quantifying respiratory status. While many methods have been proposed, few have demonstrated the ability to monitor for multiple signs of respiratory distress, and similarly few have demonstrated sufficient robustness for clinical use. The latter issue holds especially true for inertial sensors tracking the motions of breathing, as they also detect gross body motions as noise that must be filtered. Still, the compact designs of modern inertial sensors are an ideal form factor for both clinical and ambulatory care settings.

To address this set of research questions, we designed the Analysis of Respiratory Kinematics (ARK) system to quantify clinically-relevant metrics of respiratory health with the goal of predicting negative patient outcomes. ARK employs multiple motion sensors with inertial measurement cores at key anatomical locations on the chest and torso, tracking motion along the chest wall. ARK applies novel methods for analysis of respiratory data based on robust signal processing techniques previously verified in other clinical arenas, extracting known markers like a noise-filtered respiratory rate time series and resulting rate variability, as well as novel quantitative metrics for warning signs like the recruitment of accessory muscles and respiratory alternans. To facilitate design of clinically useful metrics, ARK has been deployed in a medical exercise lab, emergency department, and hospital wards and intensive care units. Preliminary predictive analysis shows that ARK produces metrics that can discriminate between patient outcomes and have statistical value for outcome modeling. This work represents one of the first applications of robust, multiparametric respiratory information derived from inertial sensors to generalized clinical outcome modeling and prediction, validating the potential for inertial respiratory sensing even in noisy conditions.

Dedication

In memory of my grandfather, Dr. Herschel Gore, Jr. (CLAS '51, SMD '55),
and a friend for anyone and everyone, Zachary Brunt.

Acknowledgements

I would like to thank my advisor, Prof. Ronald Williams, for giving me a chance with this project, for constantly offering assistance and wisdom (often in the same breath), and for strongly advocating for me for the past four years. I have learned a great deal from his guidance as well as his patient answers to my seemingly endless lists of questions. Beyond academics, Prof. Williams has been a role model for how to conduct oneself professionally while still maintaining a healthy sense of humor and perspective. I will be using his lessons on how to live healthily, happily, and productively for the rest of my life.

Similarly, I would like to thank Dr. Shrirang Gadrey for his helpful mentorship since our first summer together. I am a better researcher because of his continual efforts to bring me up to speed and his openness to sharing his domain expertise and clinical experience. Additionally, I would never recommend breaking a foot, but if you do, negotiating an emergency room is much easier with a caring doctor like Dr. Gadrey by your side. I would like to thank our ARK collaborators; our interdisciplinary team has benefited immeasurably from the efforts and advice of Dr. Randall Moorman, Prof. Sarah Ratcliffe, and Camille Hochheimer. I have immensely enjoyed the opportunity to hone my own skills under their guidance. Similarly, I would like to recognize the efforts of the other student researchers who have worked on ARK: Sarah Innis, Julia Shanno, Swet Patel, Brendan McNamara, and Arjun Deopujari. I am lucky to have worked with such a talented group.

I have been fortunate to have had many mentors beyond our research team that have shaped my understanding of engineering and its role in the world over my last eight years at U.Va. I would like to thank Prof. Harry Powell, Prof. Joanne Bechta Dugan, Prof. Todd Delong, Prof. Stephen Wilson, Melissa Goldman, Prof. John Lach, and Prof. Luke Dahl, among many others, for routinely taking time to listen and help with my academics and research; I was always headed in the right direction after our conversations. Melissa and Prof. Lach were instrumental in opening my mind to the idea of going to graduate school, and both helped build my confidence that I could meaningfully succeed here. I would like to thank the Massie Family and the Jefferson Scholars Foundation for their generous support during my time at the University. Ben Skipper and Katie Cowen at the Foundation have helped me define my own terms of success and work to meet them, and I am thankful for their continual assistance and encouragement. I would not be here without these mentors' efforts, and I have tried to pay it forward when I could.

Finally, I'd like to thank my friends, family, teachers, and professors for supporting me throughout my educational career, and especially those who have made my health a priority in their lives.

This document contains explanations and figures drawn from the following paper, which I authored with the ARK team:
W. B. Ashe, S. E. Innis, J. N. Shanno, C. J. Hochheimer, R. D. Williams, S. J. Ratcliffe, J. R. Moorman, and S. M. Gadrey, “Analysis of respiratory kinematics: a method to characterize breaths from motion signals,” Physiological Measurement, vol. 43, no. 1, p. 015007, Jan. 2022. DOI:10.1088/1361-6579/ac4d1a

This analysis used code from or based on the following sources:

- Matlab [1]
- R [2], RStudio [3], *blandr* [4], and *ggplot2* [5]
- MPU-9250 IMU I2C Library from the U.Va. Libertas Satellite Software and Avionics Team

I’d like to thank the following members of the U.Va. community for their contributions and technical support on this project:

- My committee members, Prof. Aidong Zhang, Prof. Dugan, Dr. Moorman, and Dr. Gadrey
- Dr. Arthur L. Weltman, Lisa Farr, and the U.Va. Exercise Physiology Core Lab team (Olivia Hazelrigg, Nicole Huebner, Courtney Paxton and Jackie Dosik)
- Dr. Mark Sochor, Ashley Simpson, and Anjali Kapil of the U.Va. Emergency Medicine Research Office
- Dr. Jeffrey Holmes, Dr. Sochor, and Julie Radlinski of the U.Va. Center for Engineering in Medicine
- Melissa Goldman and Trevor Kemp of the U.Va. School of Architecture Fabrication Lab
- U.Va. Libertas Satellite Software and Avionics Team

This work was graciously supported by grants from UVA’s Division of General, Geriatric, Palliative and Hospital Medicine, UVA’s Center for Engineering in Medicine, and the Ivy Foundation.

Contents

List of Figures

List of Tables

Abbreviations and Symbols

ARK	Analysis of Respiratory Kinematics
AI	Artificial Intelligence
BLE	Bluetooth Low Energy
bpm	Breaths per Minute
CNN	Convolutional Neural Network
COPD	Chronic Obstructive Pulmonary Disease
CPAP	Continuous Positive Airway Pressure
CWT	Continuous Wavelet Transform
ECG	Electrocardiogram
ED	Emergency Department
EMG	Electromyography
EPCL	Exercise Physiology Core Laboratory
FEF	Forced Expiratory Flow
FEV1	Forced Expiratory Volume in One Second
FFT	Fast Fourier Transform
FVC	Forced Vital Capacity
I:E ratio	Inspiratory to Expiratory Time Ratio
IMU	Inertial Measurement Unit
ICU	Intensive Care Unit
MBAN	Medical Body Area Network
MEMS	Microelectromechanical System
MMR	MbientLab's Metamotion R Sensor
MRI	Magnetic Resonance Imaging
OEP	Optoelectronic Plethysmography
PCA	Principal Component Analysis
PEF	Peak Expiratory Flow
PPG	Photoplethysmography

RA	Respiratory Alternans
RAM	Recruitment of Accessory Muscles
RIP	Respiratory Inductive Plethysmography
RMS	Root Mean Square
RSA	Resting Sinus Arrhythmia
SCM	Sternoclaidomastoid
SOC	System-on-a-Chip
SPWVD	Smoothed Pseudo-Wigner-Ville Distribution
STFT	Short-time Fourier transform
USB	Universal Serial Bus
V_T	Tidal Volume

Chapter 1

Introduction

Pulmonary function testing is indispensable for understanding healthy respiratory function and ensuring positive patient outcomes for respiratory illnesses [6]. Traditional methods are well-designed and diverse, measuring relevant quantities like breathing rate, flow rate and tidal volume, expiration force, and blood oxygen saturation. The techniques are generally non-invasive, monitoring via a mask, breathing piece, finger clip, elastic chest band, or similar device. Lab testing is extremely effective at quantifying the array of metrics and has been used to study diseases like chronic obstructive pulmonary disease (COPD), sleep apnea [7], asthma, COVID-19 [8], sepsis [9], and many more. Clinicians, however, often do not have access to the wealth of known respiratory indicators due to cost efficiency, physical feasibility, patient comfort, or calibration-dependent accuracy of conventional monitoring equipment [10]. Clinical testing faces difficult constraints, requiring solutions that are designed specifically for hospital environments. Advances in clinical respiratory monitoring device are needed to close the translational research gap between known signs of respiratory distress and the feasibility of monitoring methods for a clinical environment.

While it is self-evident that improving patient care and outcomes is worthy of study, the specific importance of information about respiratory condition is less obvious. Current clinical analyses use oxygen saturation and respiratory rates as primary vital signs representing pulmonary functioning. This standard exists even though dips in oxygen saturation arrive too late for preventative care and manual respiratory rate counts are known to be inaccurate [11] [12]. Thus, respiratory condition is often informed by a qualitative assessment, looking for specific physiological signs like the recruitment of upper chest and neck muscles to provide additional respiratory effort [13] [14] [15]. Charting such information for ongoing monitoring is difficult, due to the lack of exact descriptors (e.g. "looked strained", "wheezing", "heavy accessory muscle use") [16]. Worse, known markers can be missed by less experienced physicians [17], and terminology can be inconsistently applied [18]. The problem is similar to blood pressure monitoring, where pulse strength was described as "vigorous" or "weak" before the invention of the sphygmomanometer. The key intuition was to use a common scale of pressure at a known position on the body to describe the strength of the pulse across measurement times and subjects.

A variety of quantitative methods have been proposed in recent years, using new and improved technologies to extend the state of the art and deliver better respiratory monitoring in clinical environments. Many methods for respiratory rate specifically have been explored, with some yielding highly-resolved tracking of rate as a time-series [19] [10]. Others have addressed specific use cases, like non-invasive methods for sleep monitoring [20] [21]. While each of these methods quantify respiratory functioning in ways not currently used clinically, they generally fall short of full applicability due to unreliability or limited information gain. Systems using inertial sensing [22] [23] [8] specifically show promise with lightweight form factors, improvements in cost-efficacy, and ability to measure kinematic warning signs like accessory muscle recruitment. Still, inertial sensors are commonly written off as being prone to noise from body motion and the local environment [10]. Methods using inertial signals for clinical respiratory monitoring must deal with this fundamental challenge to be considered viable for deployment.

To address these issues, we present the Analysis of Respiratory Kinematics (ARK) system, comprising a larger number of sensors than previously deployed and associated synchronization and big data algorithms that enable robust monitoring. The goal of this work is to demonstrate the ability, efficacy, and reliability of inertial sensors both in lab conditions and noisy clinical theaters. These analyses support the idea that the unobtrusive motion sensors offer the best balance between data accuracy, types of data produced, robustness, and patient comfort for clinical applications. The remainder of this document is organized as follows:

- Chapter 2 details relevant background information, previous examinations of respiratory kinematics, and employed technologies and algorithms.
- Chapter 3 describes the design of two ARK prototypes and the process of collecting data sets with each.
- Chapter 4 introduces the Hilbert transform and Berger method for tracking the time-varying frequency of a non-stationary series like a respiratory signal.
- Chapter 5 builds on the previous chapter with methods for robust determination of a single respiratory rate from a group of respiratory rate estimates.
- Chapter 6 extends the previous methods to derive multi-modal respiratory metrics using the reliability information from the respiratory rate derivation.
- Chapters 7 and 8 offer discussion of the motivating and enabling technologies, areas for future work, and some concluding thoughts.

Chapter 2

Background

Research into respiratory kinematics and typical patterns spans over fifty years, seeing state-of-the-art solutions move from chest-bound strings and pulleys to motion-capture systems with dozens of optical markers. An impressively wide array of devices and systems have been tested, but none have been widely adopted beyond small research studies. In the first half of this chapter, methods and devices for pulmonary function testing and specifically respiratory kinematics are reviewed; the latter half focuses on computational methods for extracting relevant information from different sensing modalities, modeling the motion sensors used for the present study, and handling respiratory data using a time series approach.

2.1 Respiratory Mechanics and Pulmonary Function Testing

The respiratory system consists of the lungs, breathing airways, actuating muscles, and skeletal support and is responsible for ventilation (physical flow of air) as well as respiration (cellular exchange of gases). For a typical breath, the autonomic nervous system activates the primary breathing muscles, expanding the abdominal cavity and creating a pressure gradient in the lungs that draws air through the connecting airways. Gases exchange in tiny air sacs known as alveoli, transferring oxygen into the bloodstream and carbon dioxide out. Finally, the primary breathing muscles naturally contract, reversing the pressure in the lungs and driving air out.

The primary muscle group consists of the diaphragm in the abdomen and the intercostals surrounding the rib cage. Together, the two groups work to manipulate the hinging rib cage and the assortment of organs in the abdominal cavity, making room for air to enter the chest. In the event that primary muscles are overloaded, secondary muscles kick in to provide additional force to move the chest. The bilateral scalene and sternoclaidomastoid (SCM) muscles run along the front and side of the neck to the shoulders and sternum; these secondary muscles lift the upper chest, allowing more air to flow. This pattern is common for high-exertion situations, but for a stationary patient it indicates that the patient is having trouble maintaining minimal respiratory performance without exerting extra effort. Thus, accessory muscle use is a known warning sign of respiratory decompensation.

A multitude of metrics exist for monitoring the condition of the respiratory system [24] [6]. Respiratory rate is the easiest to obtain, and can be counted manually or extracted from many different physiological and kinematic signals. Peripheral blood oxygen saturation can be monitored via photoplethysmography (PPG) with a fingertip pulse oximeter. Characterized as vital signs and closely monitored during respiratory distress cases, extreme measurements (high rate, low saturation) are telling of poor outcomes. Still, these metrics do not fully predict the outcome of all such patients, or arrive too late for preventative care [25].

Other metrics can complete the picture of respiratory health [6]. Spirometry measures flow rate and volumes [26] [27]; most commonly, the total volume of a regular breath, or the *tidal* volume is measured. A related measure is the same test over a full forced breath, giving the forced vital capacity (FVC) (see Figure 2.1) [24]. Similarly, analysis of the individual phases of the FVC gives forced inspiratory capacity and expiratory capacity. Flow volume analysis looks at flow as a function of volume as opposed to time. On a forced breath, these data yield forced expiratory flow (FEF) landmarks relative to current lung volume, peak

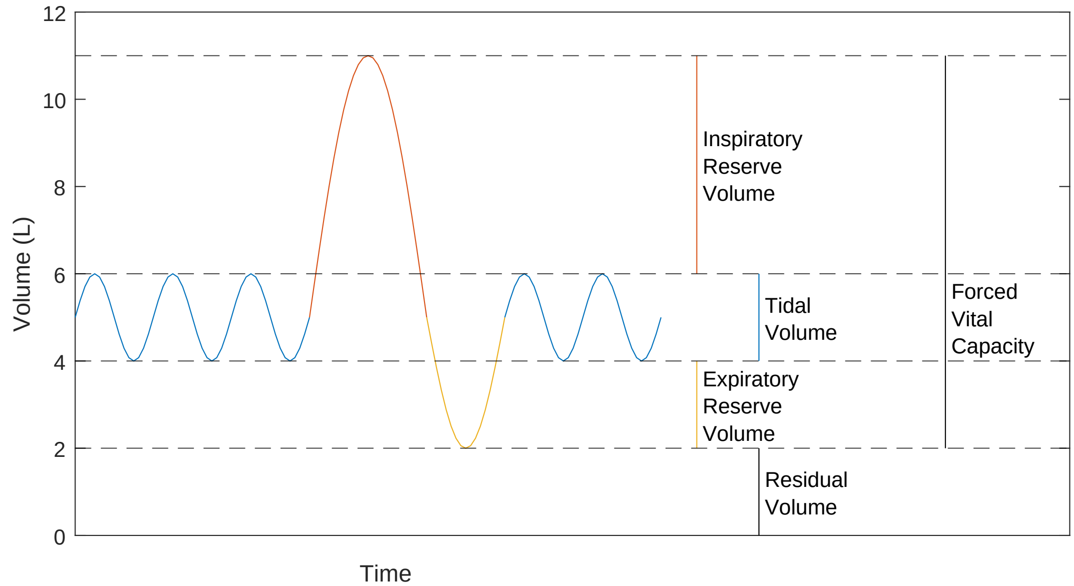


Figure 2.1: Breakdown of lung volumes as portions of the total lung capacity as determined by spirometry.

expiratory flow (PEF), flow volume loops, and the resulting maximal expiratory flow volume (see Figure 2.2). PEF can also be measured using a handheld peak flow meter and a forced maximal breath. The one-second forced expiratory volume (FEV1) is a straightforward extension of the flow volume loops. Flow over continuous, relaxed breathing yields tidal volume (V_T).

Spirometry can be combined with either nitrogen washing or gas dilution to find the residual volume and the total lung capacity [24]. For nitrogen washing, a subject breathes an oxygen supply to wash out the nitrogen in the residual volume, which is collected by a downstream sensor. The resulting total volume of nitrogen is proportional to the residual volume. For dilution, a subject breathes a mixture of air and a tracer (e.g. helium), which will equalize between the original tracer reserve volume and the subject's residual volume.

Diffusion capacity and capnography are used in conjunction with peripheral blood oxygen saturation to assess the gas transfer functioning of the lungs [24]. Diffusion capacity measures the lungs' ability to uptake carbon monoxide from a gas mixture, testing the ability to diffuse gases like oxygen and carbon dioxide across cellular membranes. Conversely, capnography studies the exhaled carbon dioxide in a subject's breath [27]. Serving as a basic function test,

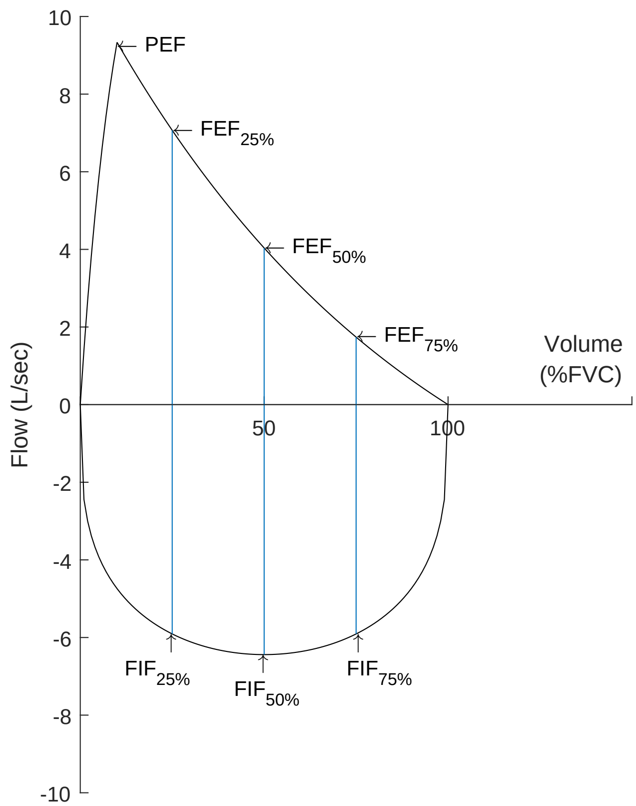


Figure 2.2: Example of a flow-volume graph. Volume represents change in volume in the spirometer (i.e. positive is amount exhaled) starting from fully inhaled breath at the origin.

capnography is commonly monitored in critical care units, operating rooms, and emergency departments.

Cardiopulmonary exertion tests study the efficiency of aerobic energy production through monitored, strenuous exercise [24] [27]. By combining synchronized electrocardiogram (ECG), respiratory rate, oxygen consumption, and carbon dioxide production signals, the analysis finds the maximum aerobic work per unit time at the peak level of exertion. The most common result, the maximum consumed oxygen per unit time ($\dot{V}O_2$ max), is a measure of the overall fitness of the subject.

Though each of these metrics provides useful information, they are all either difficult to track over time on a hospital floor or only narrowly useful [6] [28]. Spirometry and other sealed mask techniques are not feasible for long-term monitoring, as the mask could be unnecessarily uncomfortable or even aggravate the condition of the patient and impact the accuracy of recorded data [29] [30] [31]. Techniques could reasonably be applied at intake, but hourly tests or continuous monitoring would be infeasible due to the complex setup and calibration. PEF is useful for at-home asthma monitoring and can have implications for COPD, but it has not been applied broadly to other respiratory illnesses [24]. Cardiopulmonary exertion tests are not possible due to the involved trial setup and procedures; tests requiring exertion would be impractical for many sick patients. Although these established methods accurately describe pulmonary function, they leave open questions for how best to apply the knowledge in a clinical setting.

2.2 Methods for Sensing Respiratory Kinematics

Respiratory kinematics have been quantified using a variety of methods, including magnetic displacement, respiratory inductive plethysmography (RIP), magnetic resonance line scanning, optoelectronic plethysmography (OEP), fiber optics, computer vision, skin strain, and inertial sensing. The following represents a sample of research in each domain, with a focus on understanding how sensing modalities have evolved into the current state-of-the-art and how their physical design and output metrics have defined their use.

2.2.1 Magnetic Displacement

Konno and Mead [32] introduced one of the earliest magnetic displacement techniques for kinematic quantification in 1967. The system used a pair of strings, attached at the sternum and above the navel, attached to pullies. The weights on the strings partially consisted of a electromagnetic core housed inside a linear differential transducer. The resulting signals were amplified and written to tape with a recorder, for later visual review. The data representation technique is notable: rather than attempting direct numerical analysis, the synchronized two-dimensional data are displayed on a two-dimensional plane, with the line of connected points representing relative motions through time between the abdominal and chest walls. Recording data for later reviewing without tape playback required a polaroid camera. The key assumption behind the two-signal paradigm is that the movement of the chest, in many breathing maneuvers, has a single degree of freedom throughout each of the chest and abdominal regions.

In place of the strings and transducers, Hixon, Goldman, and Mead [33] proposed on-body pairs of generating and receiving magnetic coils in 1973; generator coils were attached at the sternum and one in the middle of the abdomen just above the navel, while sensing magnetometer coils were placed in similar posterior positions. Again, the technique produced 2-D relative motion diagrams, relying upon the two signal paradigm. Watson and Hixon [34] used the second method to explore respiratory kinematics in opera singers. They recorded a variety of breathing motions with their relative diagram technique across four trained opera singers. Vocal techniques included conversation, varying reading styles, and multiple paces of singing. The same testing apparatus was used to study the motions of female classical singers, with similar results [35].

2.2.2 Respiratory Inductive Plethysmography (RIP)

Like the above magnetic displacement methods, RIP uses a pair of sensing devices around the chest and abdomen. The technique senses kinematics with two elastic bands, with multiple coils of insulated wire sewn into the bands. The inductance of the coil varies with expansion, which can be detected in a sensing resonant circuit. The resulting waveforms move proportionally to the circumference of the band, producing similar displacement signals to the MD methods. [36]

First introduced by Cohn et al. in 1977, the device was quickly recognized as a monitoring tool with great potential [37] [38]. The technique has been used for exercise monitoring with lab and portable versions [39] [40]. It has also been used to quantify negative respiratory events like Respiratory Alternans [41] and vaso-occlusive crisis in sickle cell patients [42]. The devices excels for exercise applications given that the bands measure the expansion of circumferences without the motion of the body, so the signals are comparable despite occurring under stressed conditions. The technique cites Konno and Mead's two compartment model. Tobin et al. found that the technique allowed for accurate monitoring of pulmonary disease despite the observation of a third degree of freedom in the chest wall [43].

2.2.3 Magentic Resonance Scanning

In 1986, after years of research into the two-compartment model, Smith and Mead [44] detailed a third degree of freedom in the chest wall. The resulting three component model still includes the original two components, movements in the abdominal wall and rib cage. It adds a deformation component relating to posture and the positioning of the spine and pelvis. Using these three independent variables, gross motions of the chest can be quantified adequately. The authors leave open the possibility that motions of specific points of the chest wall or the attached skin moving with numerous additional degrees of freedom. This collective understanding led to a number of developments to study the non-simple behaviors of motion spread across the chest wall.

The first such development was the application of magnetic resonance imaging (MRI) to respiratory kinematic analysis. As of 1992, MRI had been successfully applied to image the lungs for to detect abnormalities in lung composition. Motion artifacts, however, proved to complicate the acquisition process for the quasi-static images of a moving system. Korin et al. [45] used a line scanning technique to observe chest composition along a single axis over time. The technique was used in multiple dimensions to quantify the velocity of the liver and spleen during normal breathing motions.

With improvements to MRI scanning techniques, Kondo et al. [46] [47] extended the technique by repeatedly looking at three orthogonal planes aligned with the lungs instead of lines along a single dimensions. They found that processing on MRIs could extract accurate lung volumes and cross-sectional areas. They also examined motion of the rib cage and found that the organization of both rib cage and abdominal motion was non-simple; measurements in different locations did not track together.

Kiryu et al. [48] used MRI techniques to observe the synchrony and relative contribution of the left and right halves of the chest in supine, prone, and each side positions. They found that the left and the right hemidiaphragmatic motions were synchronized except when the subject was lying on their side. Motions were found to differ in shape magnitude across postures and breathing phase.

2.2.4 Optoelectronic Plethysmography (OEP)

Another response to the complex motion of the chest was the development of OEP. The technique was detailed by Ferrigno et al. [49] in 1994, where the authors constructed a mesh model of the chest by tracking optical relectors positioned in a grid along the front and back of the torso. Combined with a standard set of motion capture cameras, the authors showed that the motion of specific subsections of the chest correlates with the contribution of volume from that region of the lungs. They divided the chest into left and right sides, observing three vertical compartments. The technique was quickly improved by Cala et al. [50] working with a similar team.

OEP has been applied to quantification problems for multiple respiratory illnesses with success. Romagnoli et al. [51] showed that the non-invasive monitoring could be used to study the chest movements and muscle pressures in ankylosing spondylitis patients. Similarly, Binazzi et al. [52] applied the technique in a population of COPD patients; they found that reading aloud, singing, and high-effort whispering altered the kinematics of the chest, exacerbating differences in breathing pattern and muscle recruitment.

This technique is considered state-of-the-art for evaluating breathing mechanics and correlating them with measurements of lung volume [53]. From the resulting motion data, a number of quantities can be derived, notably: respiratory rate, tidal volume, ventilation volume per minute, peak inspiratory flow, peak expiratory flow, tidal volume to respiratory rate ratio, inspiratory to expiratory time ratio, compartment phase relations, and forced flow landmarks.

2.2.5 Fiber Optics

Fiber optic sensors have also been applied for vital sign monitoring, using a scattering component known as a fiber Bragg grating [54]. The fiber Bragg grating is a periodic change in the refractive index of the transmitting material which causes a reflection for the so-called Bragg wavelength while other frequencies pass without attenuation. The key component to the sensing technology is a linear change in the wavelength due to temperature changes or physical deformations. Analyzing the spectrum of the transmitted signal over time yields the shifts in Bragg wavelength, allowing for reconstruction of the deformation signal. Further, because other frequencies pass unhindered, multiple gratings with different Bragg wavelengths can be constructed, allowing for multiple sensors to be chained together with a single optical source and receiver pair. When fixed to the chest with a circumferential band similar to RIP, the deformation signal corresponds to the expansion and contraction of the chest, allowing for accurate reconstruction of the respiratory kinematic effort signal.

2.2.6 Mobile Platforms

Researchers have also been able to extract respiratory rate from a video of respiratory motions from a smartphone camera [55]. Researchers extracted the necessary oscillating signals from the varying intensity values of pixels located over the chest wall, producing a similar signal to fiber optic and RIP solutions. The researchers extracted an instantaneous rate using smoothed pseudo-Wigner-Ville distribution (SPWVD), a time-frequency technique with more parameters than a straightforward short-time Fourier transform (STFT) that allow for independent smoothing in the time and frequency domains. The approach is well-reasoned and particularly effective in home and ambulatory care settings, where specialized equipment is not available but smartphones are. For clinical settings, however, computer vision brings practical problems of finding a reliable position and building detecting algorithms for complex features. Additionally, a camera in a hospital environment becomes highly invasive and would entangle all manners of privacy concerns. While the techniques show promise, the method is not quite fit for clinical venues.

The availability of mobile devices has also driven other creative solutions, using inertial sensors and microphones in phones [56], sensor arrays in smart watches [57], and respiration sounds from earbuds [58]. Similarly to the computer vision system, these solutions are inventive and useful for non-clinical settings where requirements of robustness are lower and the sensors are plentiful. That said, they are not the most effective prior art for a clinical arena.

2.2.7 Wearables

Inertial Sensing

Numerous studies have examined the relationship between inertial signals and the true breathing rate, at a number of locations and the potential for use in a clinical environment [59] [10] [60] [61]. Inertial sensing is performed with an inertial measurement unit (IMU), with the ability to measure some combination of acceleration, angular rotation, and magnetic field. The idea has been extended to ubiquitous sensing platforms like smartphones [57], enabling advances in inertial processing to even be applied in home environments. Multiple studies have reviewed the respiratory rates derived from various locations, finding small or no differences between rates produced [62] [63]. Methods for deriving rate include peak/cycle counting, spectrum analysis, and real-time statistical analysis to iteratively track breathing direction. Methods explained below serve as prior art for analyses concerning clinical application or the use of sensor networks to improve results.

Much of the early application work concerns obvious applications of inertial sensors to open problems in medicine. Sleep apnea has been a prime candidate, given the motion component of breathing and the importance of reducing patient discomfort to retaining good data. Apnea systems [21] [64] [20] [65] have been validated with around 90% accuracy, either by tracking the motion of the chest and the lack thereof during apnea, or by sensing the sounds associated with breathing from the base of the neck or directly on the septum. The motion component of apnea appears clearly on acceleration recordings, allowing for robust extractions of apneic intervals. Another application is respiratory gating for imaging [66] [60], where the real-time data from an accelerometer are used to predict intervals of low motion, reducing motion noise effects during each recording window. While these are interesting applications demonstrating the relevance of kinematic data, they do not attempt to fill the void of clinical-grade devices for assessing generalized respiratory risk. Earlier work also included multiple single-sensor measurement systems for respiratory rate: [67], [68], [69], among others.

With the advent of cheap sensors and radio/processor combinations, body sensor networks supporting multiple nodes were more easily realized, spurring research into larger configurations. [62] used a two-accelerometer device to observe waveforms from a pair of sensors on various configurations, finding good agreement between bilateral signals. They also examined breathing signal differences between the front and back of the torso, finding that subjects on their side had very little respiratory motion on their back. [59] moved to two pairs, on the upper and lower halves and front and back of the torso. They used the back sensors as references, with upper and lower breathing magnitudes computed as the difference between front and back sensors. They noted stronger breathing signals in the lower chest compartment, though without focus on the clinical connection of relative compartment utilization in breathing pathologies. [63] also examined four locations, but at varying points along the chest (suprasternal notch, upper ribs, lower ribs, abdomen). They found that the respiratory rate can be reasonably recovered from each position with good accuracy, but they also found no significant difference in the accuracy of different postures (supine and seated) or in the dimension of interest. These results indicate that inertial networks are well-suited to the problem of respiratory monitoring, as all locations on the chest wall reliably move with respiration.

Further work has gone towards exploring the potential of inertial sensors and sensor networks to produce more useful metrics and more relevant clinical analyses. In 2017, two groups presented work detailing new inertial respiratory monitoring systems for respiratory kinematics with useful analytical paradigms. Cesareo et al. introduced a three-receiver system in [22] and [70] which observed rib cage, abdomen, and reference motions. The receivers use a sensor fusion algorithm to convert the inertial data into a rotational quaternion, a 4-tuple representing the axis and angle information of the rotation needed to recreate the measured orientation from a base frame of reference. The technique extracted the quaternion component containing the largest variation from each chest sensor and processed the result with standard signal processing to extract the underlying rate. Cesareo et al. [71] extended their own technique by extracting landmarks corresponding to the start and end of exhalation. Thus, their current system produces rate and inspiratory:expiratory (I:E) ratio, though the relative error on the inspiratory and expiratory times limits the accuracy of the resulting I:E ratio.

In parallel, Gaidhani et al. [23] deployed a two-receiver system, with a chest location at the suprasternal notch and a reference node at the equivalent posterior position. Rather than applying single-sensor analysis, the two receivers are viewed as end-effectors in a unified robotic measurement paradigm. The kinematics can be explicitly connected with the Denavit-Hartenberg convention, enabling the measured kinematics to be converted to measurements corresponding to displacement of the chest wall. The researchers examined normal breathing and motion changes during a coughing event; they observed that the calculated angle between sensors in the saggital plane along the body corresponds well to breathing motion events when both euler angles detect the pattern of breathing. While this is useful, not all breathing patterns or positions will generate appropriate data on both sensors to make this effective.

A team of researchers based out of Northwestern University have been examining multi-function inertial sensing with a single-receiver system at the suprasternal notch [72] with a more application-oriented approach. The sensor has a very reduced form factor, fitting a processor, radio, accelerometer, and other necessary components to a thick, band-aid-sized profile. Their fabrication technique can produce flexible sensors such that the final product is comfortable and highly inconspicuous, allowing for long-term monitoring. Their studies examined respiratory markers like respiratory rate estimation and cough event description via feature extraction. With combined gait and heart rate (R-R interval) monitoring, the team developed classifiers for inpatient and at-home monitoring of patients known or suspected to have COVID-19 [8] [73]. The classifier applied digital signal processing analysis to find the four sets of features and a convolutional neural network (CNN) to classify the feature data. The model was highly effective at differentiating inpatient cases and decently effective with at-home monitoring of less-severe cases. That being said, the use of a CNN raises questions of reliability and interpretability, and doesn't necessarily generalize to the phenotypes of other breathing ailments.

Methods have begun to meaningfully move beyond average rate aquisition as a measure of respiratory rate activity. [19] used the projection of the gravitational vector as read by an accelerometer in the earth-frame horizontal plane as an oscillating two-dimensional signal to derive a breathing rate. They used a recursive, constrained principal component analysis (PCA) to recover the direction of maximum variance in the computed plane. The signal along this direction of variance oscillates in phase with breathing, allowing for accurate means to be recorded in a variety of postures. The method also details a quality index equal to the desired number of points to be kept, allowing for a parameterized accuracy. This method uses creative processing to achieve a clean derivation of an effort signal, which would be useful for inspiration/expiration interval processing and rate variability analyses. Still, it makes no use of the depth of signal and the associated potential for predictive power. Other techniques have been proposed for instantaneous respiratory rate from an inertial sensor, including [62] which took breath intervals and created a rectangular instantaneous rate signal from their reciprocals. [69] used an acceleromter, but tracked the oscillations in breath sound frequency, rather than the respiratory kinematics. Other instantaneous rate solutions have been proposed [57] [74] [75] [76] [55] [77] [78]; these methods typically use a cleaner oscillation source like ECG or PPG, making the extraction of a single frequency easier. To bring together the methods, a successful solution should draw on the knowledge of sensor design, placement, and error models as well as the instantaneous rate literature to robustly derive a respiratory rate time series and other kinematic metrics.

Strain Sensing

In 2019, Chu et al. [26] developed a method for monitoring the expansion of the chest wall with a pair of strain sensors attached at the lower ribs and the abdomen. The resulting signals were processed for respiratory rate and shown to be closely correlated spontaneous volume from a spirometry system. The correlations improved when modified from a power regression of either single sensor to a multivariate linear regression of both signals combined. Models were used to calculate exhalation volume for each breath with good agreement and low bias, but training requires a personalized calibration. To facilitate time synchronization, all multiple-sensor measurements were conducted on the same acquisition device with synchronized sampling. Additionally, one subject was asked to breathe normally for an extended duration with a single, bluetooth-enabled ribcage sensor, with recordings taken in two minute segments roughly every fifteen minutes.

2.3 Clinical Methods for Respiratory Monitoring

Of the tests described above, the most commonly deployed metrics are respiratory rate and blood oxygen saturation. The prevalent method for oxygen saturation is PPG, which accurately characterizes oxygen content by measuring skin absorption at one of many points on the body. While models for different conditions and different physiological factors could still improve results, generalized improvements to the method would most likely require invasive procedures to quantify the actual amount of oxygen in arteries. Regardless, respiratory kinematics does not offer a simple paradigm for measuring saturation. Respiratory rate, however, is widely recorded and used, and moreover the methods used are less accurate, more cumbersome, or more limited in the captured modalities of respiratory information.

Original methods for measuring respiratory rate are rooted in counting methods, originally using stopwatches and manual counts [79]. While this provides a reasonable estimate of average rate, accuracy is based on the training of the recorder and no variability information is captured.

The advent of electronic instrumentation led to numerous developments in sensor technology, including ECG, PPG, RIP, and skin strain sensing, among others. All three have been shown to capture respiratory rate information with the oscillation of a signal of interest. Electrocardiograms can detect changes in rate due to coupling between the heart and respiratory rates [80]. Specifically, the Respiratory Sinus Arrhythmia (RSA) is a well-studied phenomenon where the interval between spiking ECG R-waves (RR interval) oscillates at the rate of breathing. Plotting and smoothing the series of RR intervals recovers the respiratory signal.

In addition to RSA detection, the size of the R-wave spike on an ECG signal is also indicative of respiratory motion via the alteration of chest impedance. The electrical impedance to signal propagation varies as the chest expands and contracts for breathing motions, altering the amplitude of signals measured on the surface of the chest. The difference in attenuation between expanded and contracted chests generates proportional amplitude variation in the R-waves of the ECG signal. Extraction of the envelope of R-waves gives a signal that oscillates at the rate of breathing.

Similarly to ECG, PPG allows for indirect analysis of repiratory rate via oscillations in the original signal [74] [81]. The skin absorption readings are sufficiently accurate and well-resolved that an unfiltered plotting of the absorptions will vary with respiratory activity.

Likewise, chest impedance [76] and electromyography (EMG) [82] have both been shown to be capable of detecting respiratory rates through the changes of the electrical parameters of the chest wall.

RIP allows for direct measurement of breathing movement by increasing the inductance in an elastic loop around the chest [37] [36]. Resulting signals clearly contain rate information, which could be reconstructed to study respiratory variation. Similarly, strain signals detect the same physical expansion, albeit with a point measurement rather than circumferential [26] [54]. Both reconstruct respiratory effort signals with high accuracy.

Though the above methods demonstrate the wide extent of knowledge on how to extract breathing rate, none of them are widely employed in medical arenas. Implementation of ECG and PPG techniques will certainly become more common in medical environments due to the prevalence of the sensing modalities. It would be relatively simple to integrate the required algorithms into new versions of the devices deployed. Still, it would be difficult to convince many administrators responsible for medical equipment acquisition that the algorithms that simply derive a respiratory signal would be sufficiently useful. This drawback would hold even more strongly for the strain method, which would not be able to provide any additional analysis other than rate and rate variability information without extensive calibration. RIP improves on this by measuring multiple compartments and allowing comparison metrics, but the dual-banded form factor is non-ideal for a noisy clinical environment.

One potential solution is the use of inertial sensors. Inertial methods have seen numerous processings to extract rate information. In addition spectrum and peak-finding methods applied to the raw inertial signals, numerous pre-processing routines have been applied with notable success by isolating clear respiratory oscillations before examining the frequency content. A single inertial sensor will have between three and nine dimensions, allowing for statistical approaches that leverage duplicated data to detect the unified breathing rate. One method uses PCA on the projection of the accelerometer signal into the horizontal plane, which isolates the direction of the oscillating component of gravity [19]. Another method uses sensor fusion to derive an orientation quaternion before processing for rate information [70]. Quaternions are complex mathematical objects used to represent 3-D orientation, where the imaginary component has three parts as opposed to one; all four parts vary over time as orientation of a sensor changes. The dominant component of the resulting quaternion was identified and used and the oscillating signal of interest. Respiratory rate has also been successful by deriving heart sound time series from the vibrations of the chest wall captured with an inertial sensor [83]; the interval between the electrical impulse (R-wave) and the first heart sound (S1) is highly consistent, allowing for extraction of a measure that closely tracks R-R interval. Oscillations in the resulting S1-S1 interval series occur at the rate of breathing.

While the ability to detect rate, and specifically time variations in rate, has been demonstrated on multiple inertial platforms, only limited analysis has gone towards quantifying and studying respiratory rate variability. [84] found that higher variations in the average rate were associated with admissions to the intensive care unit (ICU), finding that the coefficient of variation percentage of readings just before ICU admission were significantly higher than previous variability. Extending the concept of multi-hour variability, [85] introduces breath-variability metrics on a short term basis, examining interval time series from yoga practioners and controls with a variety of analytical methods. They noted changes in standard deviation of breath intervals and breath interval differences, spectral density via interpolation and the Lomb-Scargle method, autoregressive modeling, wavelet transforms, and entropy methods like sample, approximate, and Renyi entropies. Their analysis suggested large potential for various measures, though with little statistical testing to explore the usefulness of each metric. Further, the analysis was performed with data from PPG; it remains to be demonstrated that inertial sensors can provide similarly high quality variability results.

Further, little work has been done to quantify other motion patterns known to be useful in a clinical setting, like the recruitment of accessory muscles (RAM) [15] or respiratory alternans (RA) [6] [86]. RA is a detectable transition in primary breathing motion contribution between abdominal and thoracoabdominal (rib cage) movement, indicating fatigue and an inability to meet respiratory demands. RA has been quantified with motion platforms like RIP, using two inductive bands to monitor the relevant contibutions of the ribs and abdomen [41] [42]. Inertial networks have not tackled this problem, likely due to the increased difficulty in interpreting acceleration signals (as opposed to position) and the necessary synchronization requirements so that the signals from different sensors can be directly compared in time. The latter becomes more difficult on large systems, as each sensor commonly has its own processor and clock.

Most notably, inertial methods have not been used to generate a respiratory rate time series that allows for consistent analysis. Additionally, no work has gone towards quantifying accessory muscle use, despite its status as an indicator of respiratory distress [15]. Inertial sensors have all the benefits of RIP with regard to the potential for multi-modal analysis, packaged in a desirable form factor that could be easily applied to patients regardless of current status. Still, more work is required to capitalize on existing methods and push lab results into a clinically usable form.

2.4 MEMS Technology for Inertial Sensing

Sensing paradigms using IMUs have gained traction in recent years due to the proliferation of highly cost-efficient sensors made using microelectromechanical systems (MEMS). MEMS devices consist of electromechanically coupled systems and the necessary hardware alongside digital controllers on wafers with very small form factors [87]. Accelerometers and gyroscopes can be designed using suspended spring-mass systems, with the mass generating a capacitive voltage proportional to the displacement experienced by the mass. Both can measure forces along three orthogonal axes, with each dimension using a separate spring-mass system. Accelerometers feature a linear sensor tuned with spring and dampening effects to capture motions in the frequency band of interest; the force applied to the sensor results in displacement of the mass, which can be reversed using the known dynamics of the system. Rate gyroscopes are more complex, leveraging a similar tuning scheme in two dimensions to measure the Coriolis effect. For a rotating body, the sensor inertial frame will experience a displacement in the direction of rotation relative to the world reference frame. If the mass is driven into an oscillation along a single dimension in the sensor frame, then the rotating force appears as an oscillating perpendicular force and resulting displacement. The rate of the perpendicular oscillation is the same as the driven dimension, but the amplitude is proportional to the angular rate experienced. Again, the true angular rate can be reconstructed using the electromechanical system dynamics. [87] provides a detailed introduction to the realization of MEMS accelerometers and gyroscopes with helpful visualizations.

The above processes allow for tiny sensor units to be constructed, with complete sensing packages having areas measuring less than $10mm^2$. Further, advancements in both low power analog and digital electronics and microfabrication methods have driven the price under \$10 a unit, often with three-axis Hall-effect magnetometers that round out the inertial sensing capabilities of the IMU. For all of the above reasons, MEMS IMUs represent the state of the art for wearable inertial sensors fit for explorations of human biomechanics.

Completing the view of the MEMS IMU is a proper error model, based on [88] and [89], to accurately quantify the inaccuracy and imprecision of these devices. While they are certainly usable enough to be considered cost-effective for many applications, the low-cost aspect of their design does appreciate into notable noise and bias. Consider the following nine-axis model for a MEMS IMU:

$$ {}^{s}\mathbf{a} = {}^{a}\mathbf{f}(\mathbf{C}_{e}^{s}(\mathbf{g} + \mathbf{a}_{motion})) + {}^{a}\mathbf{b} + {}^{a}\mathbf{v} \tag{2.1} $$

$$ {}^{s}\boldsymbol{\omega} = {}^{g}\mathbf{f}(\boldsymbol{\omega}_{true}) + {}^{g}\mathbf{b} + {}^{g}\mathbf{v} \tag{2.2} $$

$$ {}^{s}\mathbf{m} = {}^{m}\mathbf{f}(\mathbf{C}_{e}^{s}(\mathbf{h} + {}^{h}\mathbf{b})) + {}^{m}\mathbf{b} + {}^{m}\mathbf{v} \tag{2.3} $$

where

- ${}^{s}\mathbf{a}$, ${}^{s}\boldsymbol{\omega}$, and ${}^{s}\mathbf{m}$ are the sensor-frame acceleration, angular rate, and magnetic field readings,
- $\mathbf{C}_e^s$ is the current rotation matrix between the earth and sensor frames,
- $\mathbf{g}$ and $\mathbf{h}$ are the gravitational and magnetic reference vectors (magnetic field values shown here for the Rotunda at U.Va.):

$$\mathbf{g} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \mathbf{h} = \begin{bmatrix} h_x \\ 0 \\ h_z \end{bmatrix} = \begin{bmatrix} 21.5506 \\ 0 \\ -45.8755 \end{bmatrix} \mu T \tag{2.4}$$

- $\mathbf{a}_{motion}$ is the earth-frame linear acceleration experienced by the sensor,
- $\boldsymbol{\omega}_{true}$ is the true angular rate in sensor-frame,
- ${}^{h}\mathbf{b}$ is sum of local magnetic fields, resulting in an earth-frame bias,
- ${}^{a}\mathbf{b}$, ${}^{g}\mathbf{b}$, and ${}^{m}\mathbf{b}$ are sensor-frame bias vectors,
- ${}^{a}\mathbf{v}$, ${}^{g}\mathbf{v}$, and ${}^{m}\mathbf{v}$ are zero-mean, uncorrelated white noise process vectors,
- each $\mathbf{f}$ is a nonlinear function representing the total effects of imprecise dimensional scale factors, nonlinear responses across the measurement range, cross-axis sensitivity, and temperature distortions.

The above model comes with typical assumptions for reliably simplifying analysis. Each ${}^{x}\mathbf{v}$ is assumed uncorrelated with covariance matrix $\boldsymbol{\Sigma}_x = \sigma_x^2 \mathbf{I}_3$, with σ_x being the standard deviation of sensor measurement error. Biases are generally not highly volatile but can be dynamic, particularly for gyroscopes, necessitating online monitoring and correction for applications reliant on stationary inputs [90] [89]. Temperature distortions are corrected by the

MEMS device, which commonly contains an on-board thermometer for this purpose. Scale factors and cross-axis sensitivity can be modeled given calibration data, but the accuracy depends on the quality of calibration and the effects are often negligible. Non-linearity is more difficult to model and correct, but similarly it is often not necessary due to smaller influence. A simplified model can be constructed from the above, a combination of the models used in [88] and [89]:

$$ {}^{s}\mathbf{a} = \mathbf{C}_e^s(\mathbf{g} + \mathbf{a}_{motion}) + {}^{a}\mathbf{b} + {}^{a}\mathbf{v} \tag{2.5} $$

$$ {}^{s}\boldsymbol{\omega} = \boldsymbol{\omega}_{true} + {}^{g}\mathbf{b} + {}^{g}\mathbf{v} \tag{2.6} $$

$$ {}^{s}\mathbf{m} = \mathbf{C}_e^s(\mathbf{h} + {}^{h}\mathbf{b}) + {}^{m}\mathbf{b} + {}^{m}\mathbf{v} \tag{2.7} $$

The takeaway from the above is that MEMS devices perform their best when a strict calibration procedure is applied, or the data are analyzed with robust methods designed to exploit the signal-carrying qualities of the experimental design.

2.5 Time-Series Methods

Time series analysis is a comprehensive analytical toolkit, taking advantage of properties of data that are specifically measured to be changing in time to derive information relevant to the underlying system behavior. This information can be used to learn about the studied system and quantify expected behavior, which would be useful in the area of clinical respiratory sensing, as few platforms have been deployed for clinical studies. A larger goal would be to be able to predict negative events by identifying the warning signs from previous respiratory data. The motivating example in the study of generalized time series metrics is [91], which developed a massive database of thousands of algorithms that can be applied to time series data to discern patterns automatically. Dubbed ”Highly Comparative,” the researchers make the case that entire data sets can be characterized by the most useful metrics, which often capture the variation of the samples in a reduced form. The technique is incredibly useful and likely could lead to an interesting analysis of respiratory kinematics data, but the very idea is enough of a starting point for examining our data. Below is a brief overview of a variety of time-series algorithms that have some pertinence to the following study. Techniques are useful for handling both raw data and derived signals like a respiratory rate time series.

The simplest time series approach is the use of moments; moments are clear summary statistics that are easy to compute and widely applicable. For respiratory work, the average rate is frequently used [24], but it is less commonly computed as the mean of a time series. Such a respiratory rate time series also allows for higher moments to be computed and examined; variance would have clear implications for the rate series [85]. Higher moments can help describe the shape of the distribution and quantify the different effects of being near and far away from the mean. These properties could be useful for identifying unusual and relevant events in respiratory time series. That said, moments quantify their respective summary metric, but by definition do not tell the full story for rich effects of dynamic systems.

Another typical set of metrics are correlation functions, which examine the similarity in a reference signal to a time-shifted version of the same signal or a different signal altogether. The dot-product similarity procedure capture relatedness in signal shape, while the lag component helps quantify if there is a notable time delay or even a periodic component to the pattern. Motion sensors are notoriously noisy; correlation is a potential method for discerning whether a sensor is recording meaningful data, as other signals would be oscillating at similar rates. While the similarity procedure seems like a good fit for matching signals, correlation functions require selection of windows for computing lags and some interpretation function (i.e. a correlation threshold) to determine how to use the result.

Auto-regressive models offer a time-domain method of finding coefficients that predict future values as weighted sums of previous measured values [92]. The model coefficients describes correlation in the underlying process, offering a different view to sample change than a typical frequency-domain technique like the Fourier transform [93]. The technique requires that the series is stationary, with constant mean and variance, which would limit applicability to unstable breathing processes. The integrated model can extend the auto-regressive technique to non-stationary series that become stationary after taking some number of sample differences, essentially producing a new time series for analysis using the difference operator. Combined with moving average white-noise model, the ARIMA model is a comprehensive structure for describing the stochastic progression of a time series in terms of past realized values [94]. While ARIMA models can be powerful, they rely on assumptions that don't necessarily hold for highly-dynamic, complex processes, potentially limiting their application to more regularized modeling problems. Still, they are powerful alternatives to conventional frequency domain approaches when applicable.

The set of spectral methods offer a different view to information recorded in the time domain, looking at the important frequency components that determine the behavior of a signal. The coefficients derived from a Fast Fourier Transform (FFT) offer the most straightforward analysis, estimating the frequency content across a set of frequency bins. This technique is powerful and general, allowing for analysis of artistic and scientific images [95], sound and music [96], physiological recordings [97], cosmological data, and any other signal uniformly sampled in either time or space. An FFT is a frequency-specific averaging operation, providing a one-to-one transformed representation of the original data. While the FFT is an intuitive option for respiratory data and has been used extensively to quantify respiratory rate, the lack of information about the timing of prevalent frequencies means that the resulting coefficients do not necessarily track prominent frequencies. Most importantly, the component with the highest magnitude is not guaranteed to match the actual overall average respiratory rate. That said, the FFT still has applications to respiratory modeling through modification and time series analysis beyond raw respiratory data.

One potential solution is to use a time-frequency method, like the STFT [19] or the continuous wavelet transform (CWT) [74]. While they differ slightly in computation, they both generate spectral magnitudes evaluated at multiple points in time over the course of the signal. The process allows for a controlled balance between time and frequency resolutions, but techniques are fundamentally limited by the Heisenberg-Gabor uncertainty principle; increases in frequency resolution inherently decrease possible time resolution. A more ideal method would allow for the two components to be handled separately, or at least allow for optimal integration of frequency analysis in time. The SPWVD method [55] uses independent time and frequency window functions to calculate the time-frequency representation, solving some of the issues of the STFT and CWT. The computational difference is that both the STFT and CWT are invertible, while SPWVD is not; for the former, filtering is possible to isolate content of interest in the frequency domain before converting back and continuing analysis on a time-domain signal.

Multiple solutions have been presented for working specifically with the instantaneous frequency of a signal. The most popular candidate to model these non-stationary signals is the Hilbert transform, which shifts frequency components by 90^o to create a complex signal from a real-valued input [27] [98]. The resulting analytic representation oscillates around the origin of the complex plane; the complex representation allows for tracking the phase of the two-dimensional signal, which corresponds with oscillations in the original. This phase is the integral of the instantaneous frequency, though simple derivations based on discrete derivatives are often too noisy to be used [75]. An improved approach is the Hilbert-Huang Transform [99], which derives a set of basis functions (intrinsic mode functions) which are well-behaved inputs for the Hilbert transform, creating cleaner instant frequency representations for each component. These basis functions act as both a decomposition of the original signal and individual analytical objects whose non-stationary frequency content can be examined with the Hilbert transform. The synchrosqueezing transform [74] works similarly to track ridges in the time-frequency representation that correspond to mode functions in the original signal. The synchrosqueezing transform builds on either the STFT or the CWT; it provides a mathematical basis for deriving basis functions out of the time-frequency data that can be inverted back into the time domain using the underlying transform inverse. This technique is more theoretically grounded in mathematics than the Hilbert-Huang Transform, but it still relies on some assumptions about the separability of the signals in the frequency domain in order to robustly derive the correct number of basis functions.

The above methods characterize different expectations of regularity, important for understanding the typical behavior of a signal. A complementary analysis is the examination and quantification of the random behavior of the system using entropy and chaos techniques. Entropy is a quantification of expressed randomness in a random system or a sample of data derived from one. While Shannon entropy was originally designed for analyzing strict probability distributions, the idea has been extended with algorithms like Approximate Entropy [100] and Sample Entropy [101], which quantify the repetition in a time series or similar ordered data sample. Sample entropy has been used to find artifacts in heart rate (RR interval) signal regularity that have predictive implications for patient deterioration [102]. While existing entropy studies of respiratory rate-related series are limited [85], it stands to reason that such a technique could be useful for quantifying unusually regular or irregular patterns. Other methods from chaos theory could be useful as well, like Hidden Markov Models for tracking the underlying state of a system and understanding the dynamic changes in system parameters driving variations and artifacts in the original data.

Chapter 3

Overview of the Analysis of Respiratory Kinematics (ARK) System

Building on existing works, one potential solution to the problem of robust clinical monitoring is a network of more inertial sensors, with more complex processing to take advantage of the additional information. Previous kinematics studies have demonstrated efficacy of different positions, but they have done so without meaningful study of interactions between waveforms from different positions past comparisons of peak rates from a spectrum. Here, a novel network of time-synchronized inertial sensors is proposed, designed to robustly capture inertial data ready for a plethora of analytical possibilities. Two prototypes are described: a wired version used in a small laboratory study, and a wireless platform deployed in an emergency department (ED) and hospital wards. Data sets from each set of trials are introduced with some preliminary signal processing.

3.1 Design Methodology

Multiple inertial networks have been deployed with some success, though each of the designs have used different locations. Each prototype has demonstrated that those locations have value, so it follows that synchronized measurement of all locations would bring the model closer towards the OEP gold-standard. Equivalent measures from different sensors would provide fidelity and reliability, while comparisons of behavior of different regions would yield useful information about the synchronization of breathing physiology. Adding more sensors, however, is non-trivial, and increases the communication burden. This is particularly true in wireless designs, where endpoints share bandwidth with whatever other signals exist in the measurement space.

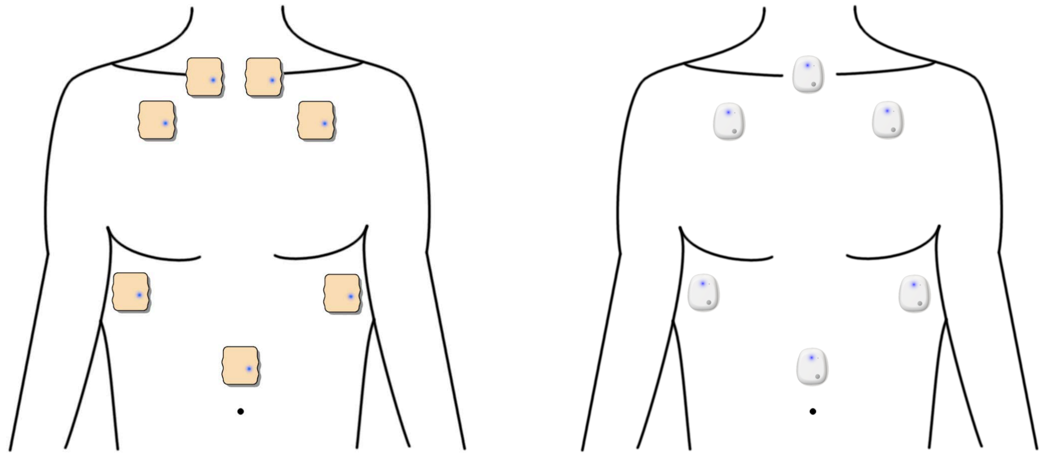
For the intertial technology, MEMS offers the necessary accuracy and precision in an enticing form factor. MEMS devices typically come with an onboard processor for managing sampling functions like output rate and range, providing a digital interface for the data acquisition process. Beyond sensing hardware, the sensor needs some combination of hardware that allows for central processing, communication, and synchronization. Synchronization is imperative for a multi-sensor network, especially in an understudied context. Post-processing for the respiratory kinematics problem would be difficult, as curve registration would depend highly on the employed breathing maneuvers. Given that the objective is to study a coordinated view of the chest, the underlying measurements must be made in comparable time scales.

Given that synchronization has tight requirements for communication, the two must be selected together. Common communication mechanisms are wired serial protocols and wireless Bluetooth radio modules. Wired solutions are generally more reliable, simpler to build, and offer extremely low latency relative to most wireless solutions, as the physical layer of communication is much simpler. Additionally, many processors offer communication hardware as a attached peripheral device, simplifying hardware design. The wires, however, present physical constraints of the system, as every node needs a connection. For a system with seven or eight nodes, this effect is appreciable.

A common alternative is Bluetooth Low Energy (BLE), a wireless communication protocol designed to be flexible to the interactions required for use cases like audio streaming, item tagging and tracking, and data acquisition platforms. Due to the popularity, there are numerous BLE-enabled processors offering a radio interface with integrated processing power on a single System-on-a-Chip (SOC), specifically designed for simple wireless devices and sensors. The integration speeds up development, but by nature of the protocol the development cycle is still more complex than a wired equivalent. Another consideration for wireless platforms is the electromagnetic noise conditions in the deployment environment, which affects reliability, latency, and synchronization quality. Even if synchronization is provided for in a SOC and its associated software library, the process could be corrupted by local noise at the time of synchronization.

Two important concerns that interact with the above are mechanisms for data handling and power management. Data from the inertial measurement device must either be reported or stored locally; the former requires higher bandwidth, while the latter requires additional hardware. If the device is wireless or otherwise power-constrained, the energy required to run the acquisition, processing, control communication, and data handling must be closely accounted for. Modern components for wireless sensors are often low-power by design, but connections like BLE also impose additional constraints to achieve their optimized performance. Care must be taken such that the custom protocols for a given application do not stress the system beyond energy capacity.

An ideal sensor for respiratory kinematics would solve the above challenges for processing, communication, and synchronization, but also provide a physical design and form factor fit for on-body measurements. The sensor should be relatively lightweight and compact. Modern design technologies like 3-D printing allow for quick, iterative prototyping of physical components with high-quality plastics, resins, and metals. Additionally, the adhesive must



(a) (b)

Figure 3.1: ARK Sensor Locations for (a) EPCL [27] and (b) ER prototypes. The back sensor for both prototypes was placed opposite the single SCM sensor at right.

be designed for skin-contact; there exist a number of options for FDA-approved adhesive mechanisms. Lastly, a minimal user interface (i.e. an LED) is useful for programming, debugging, and monitoring of device operation.

3.2 Summary of Prototypes

The ARK system is designed to address the current inadequacies of inertial methods for respiratory kinematics with respect to long-term or regular monitoring of respiration in a clinical environment. Two iterations have been made: a wired variant for lab testing, and a wireless variant for longer in-situ monitoring of emergency room patients.

The system consists of a network of sensors with inertial measurement chips spread across the chest (see Figure 3.1). The positions studied are the collection of positions in previous studies: suprasternal (base of the SCM muscles), ribs, abdomen, and reference. Modifications include an upper rib location in addition to the lower rib location, as well as expanding from single-sided to bilateral sensing for each rib location. Bilateral locations were also used for the suprasternal location, but the attachment to the body rested on the clavicle. Difficulty in leveling the sensor against the chest led to a modification to a single, centralized sensor. The reference node was placed on the back, similarly to [23]. The goal of the back-located reference node is to capture information about gross body motions that would affect all sensors, in an effort to identify and potentially correct influenced noisy data.

The next subsections detail the design decisions made for each of the two prototypes.

3.2.1 Wired System for Exercise Physiology Lab Experiments

The goal of the experiments in the U.Va. Exercise Physiology Core Laboratory (EPCL) was to simulate respiratory distress with exercise and observe the effect of peak exertion on respiratory kinematics. Exertion trials on an exercise bike usually take between thirty minutes to an hour, depending on the fitness level of the test subject. The relatively short trials were supervised, allowing for more flexibility with the requirements of the data acquisition system.

A wired design was selected to simplify the design process. The sensor was based on the MSP432E401Y processor, with a 120MHz Arm Cortex M4F and peripheral universal serial bus (USB) 2.0 PHY module acting as a USB end-point device. The inertial core was a MEMS-based MPU-9250 IMU, with three-axis accelerometer ($\pm$2g), gyroscope ($\pm 250°$/s), and magnetometer ($\pm 4800\mu$t) sampled at 100 Hz. The resulting output data were recorded with 16-bit precision for accelerometer and gyroscope, and 14-bit precision for magnetometer.

Communication over the USB connection allowed for control commands (i.e. synchronization and acquisition operation) and data transmission commands from a USB master application running on a PC. The acquistion app was built in LabView, with functions for synchronization and data collection. Synchronization was done using a one-way time-of-flight measurement, enabling the estimation of original send time for a periodic pulse train to convey a common frame of reference. Once synchronized, the sensors were started at a staggered interval, and collected data every ten milliseconds. The data were held in internal queues until the master device requested all available data. Each sample with timestamp (a running 10 μs counter) was packaged into a data packet format and sent over the wired

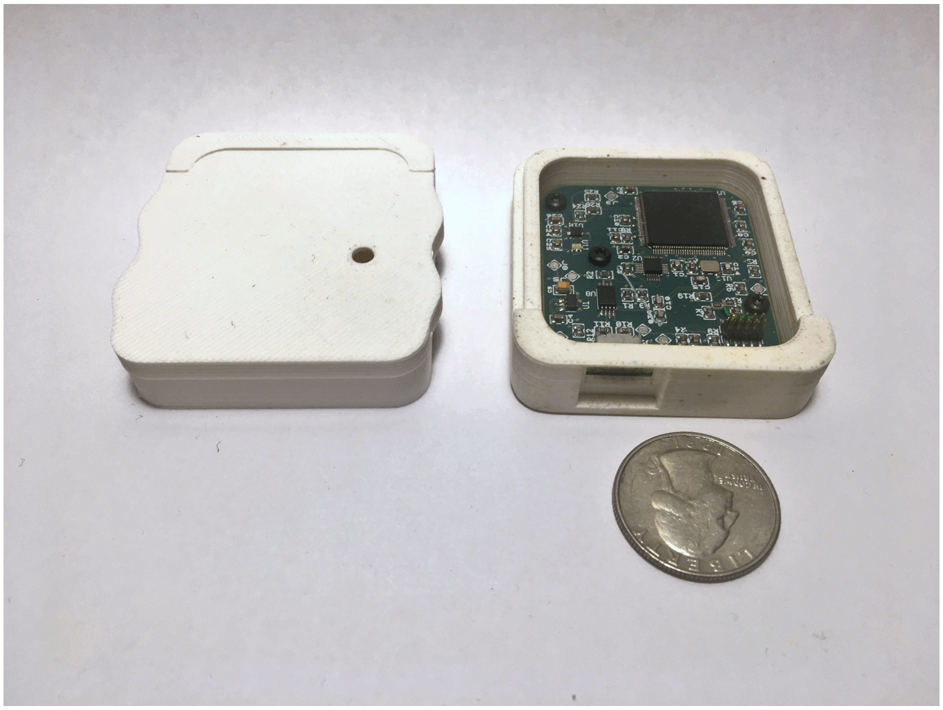


Figure 3.2: Original sensor board with sliding-top housing design, used for EPCL trials. The top locked in place, leaving an opening on the top for the LED and an opening on the side for the USB Micro B 2.0 port.

connection. Resulting synchronization was accurate to 10ms. To improve the stability of the sensor clock, a 25 MHz clock crystal was added to the board.

The hardware was routed on a custom printed circuit board and housed in a 3-D printed case (see Figure 3.2). Sensors were attached to the body with adhesive ECG pads that snapped into wells on the bottom of the case. The board included a tri-color LED and hardware driver, allowing for easy software control for indicator functions. The indicator LED was used to identify sensors and record positions before exercise trials, after the sensors had been placed.

The data acquisition app on the computer also connected to a National Instruments myRIO real-time processor, programmed with similar data acquisition code. The device sampled from the laboratory spirometer and capnograph to record flow rate and percent carbon dioxide concentration for comparison. Synchronization was done through Labview libraries to the same computer clock as the inertial sensing units. The instruments were

Figure 3.3: Second-iteration sensor: MBient Labs Metamotion R

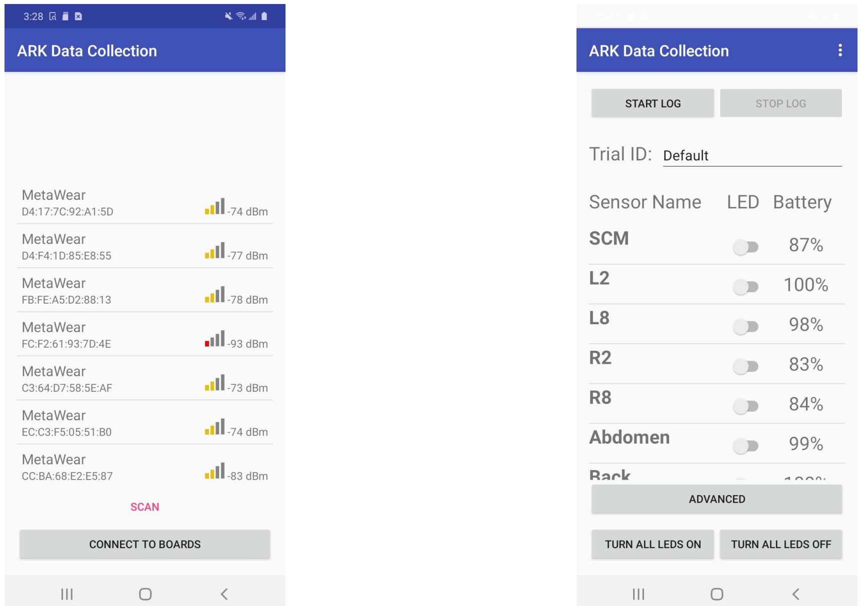
similarly sampled at 100Hz.

3.2.2 Wireless System for Deployment in Emergency Department

The prototype for the second phase of experiments was designed for the U.Va. Hospital Emergency Department. The goal of the new design was to collect longer-duration recordings from patients admitted with some amount of respiratory distress. Trials ran for approximately four hours, with multiple trials occurring at the same time. The resulting paradigm was somewhat unsupervised, though the research coordinators administering the test checked the patient and sensors regularly.

The wires in the first iteration served their purpose but were cumbersome in the lab environment. The technique would not transfer to a more chaotic emergency room, so a wireless platform using MbientLab's MetaMotion R (MMR) sensor was built (see Figure 3.3). The MMR sensor communicates via BLE to an Samsung Galaxy A10e running a custom Android collection application.

The MMR contains a BLE-processor system-on-a-chip with timestamping functionalities. Data acquisition is based off a Bosch BMI160 accelerometer/gyroscope package and a Bosch



(a) (b)

Figure 3.4: Screenshots of the (a) connection and (b) data acquisition interfaces for the custom Android application.

BMM150 magnetometer. Output values are similar, with magnetometer outputting 16-bits at an improved $\pm 1600\mu t$, but reduced to 10Hz sampling instead of the maximum 25Hz due to power concerns.

Due to the low bandwidth of BLE and the constraints of multiple BLE connections, downloading data inside the room was not feasible. The process was slow enough that switching to local storage on the MMR's 8MB NOR flash memory for later download improved duration of recording in addition to bettering reliability. At the above sampling rates, a sensor board can hold between 40 and 60 minutes of data. The acquisition of data was spaced out accordingly, with two-minute observations every half hour to ensure no data were lost.

The data collection app was built to streamline the setup interaction, starting with a one-click connection to the MMR boards (see Figure 3.4). Data acquisition was automatically configured by the app after connecting; the individual sensors are configured for output format, and timers are programmed to run the regular monitoring cycle. Starting the logging function on the phone started the timer routines on the sensor board. The sensors ran autonomously until stopped. When stopped, the phone canceled existing timers and ran one last two-minute data collection routine.

The physical design of the sensors allows for easy use in a clinical environment. Each sensor was physically labeled with it's position. Additionally, each sensor position was assigned a unique color, identified with the on-board RGB LED. Sensors could be illuminated individually or as a group from the app to check positions. The compact sensors are very lightweight, adhereing to the skin with approved adhesives provided by MBientLab.

Synchronization between the MMR and the Android device occurs at the time of download, but the drift of the MMR clocks due to clock frequency bias still presented an issue. Over the course of the four hour trial, the clocks could drift seconds apart. Additionally, the serial download process means download times were staggered, so each sensor drifted for a different period of real time. To combat the drift across all sensors, the communication events from starting the pre-programmed timers at the beginning and end were considered simultaneous across all sensors. A single communication event was measured to be accurate to 100ms, with a resulting 200ms maximum synchronization error between any two sensors. Each trial's time vector was stretched or compressed to match the average duration of the trial.

Variable recording rates further complicated the synchronization process. In the process of reviewing the data, we noted that each sensor had a different number of samples per collection, but all collections for each sensor stream had the same number of samples. This implied that the onboard IMU was generating samples at a consistent rate, very close to 100Hz. Thus, with all sensors and even different streams (i.e. accelerometer and gyroscopes on a different clock than the magnetometer), interpolation was necessary to allow for simultaneous analysis. Given that the data rates were so consistent and the time-stretching was strictly linear, the altered sensor data were assumed to be evenly spaced. Interpolation was performed by resampling a stream by a rational factor (upsample by L, filter to 1/L, downsample by M to produce L/M interpolation). All sensor data were cut to start at the latest-starting stream then cut to end at the earliest-ending stream. The result was a two-minute collection consisting of roughly 12,000 points of three-axis acceleration, angular rate, and magnetic field measurements, simultaneously sampled across all sensors.

3.3 Experiments and Preliminary Analysis

Data were collected using the technology specified below. All data analysis was carried out in MATLAB 2021a [1], with the exception of Bland-Altman analysis, which was performed in R [2] using the *blandr* package [4] with *ggplot2* [5] in RStudio [3]. Patient metadata was maintained in REDcap databases [103] [104].

3.3.1 Exercise Physiology Lab Data

For the EPCL trials, data were collected for the whole duration of the trial. The maximal exertion test was broken into two phases: exertion to lactate production threshold and max exertion. A pre-determined breathing maneuver was conducted before, between, and after the two phases. The breathing maneuver consisted of two deep breaths, followed by two minutes of unforced breathing in a supine, 30° incline position.

The physical exercise was done on an exercise bike with programmable resistance; tests were started at 60 W resistance with resistance increasing 20 W every two minutes. Subjects were outfitted with a multi-lead ECG, flow mask, and inertial device network. Exertion was measured with the Borg scale [105], rated by the user each time resistance progressed. The Borg scale is a measure of perceived exertion ranging from 6 to 20, with 6 indicating no work done (subject at rest) and 20 indicating the hardest possible effort (subject at maximal exertion). The lactate production threshold falls between 12 and 13 [106], indicating that the work is difficult enough to require anaerobic metabolism to provide the required energy. For these reasons, the lactate threshold serves as a good checkpoint between rest and max exertion conditions.

Inertial, flow, and $\%CO_2$ data for 34 healthy subjects were recorded using the above method. Five trials were unsuccessful due to inadequate sensor adherence, power failure, or unfit subjects. The physical connections of the system were points of failure, resulting in data loss for certain trials. For the analyses presented here, a set of 20 trials were used, comprising the trials where data was successfully collected for the seven chest sensors and the flow interface.

The following excerpt from our article [27] describes our synchronization process:

> Each of the eight sensors was connected via USB to a host application running on a computer. Data collection on the sensors was stagger-started over eight consecutive ticks of a common 100Hz timer on the host machine. Once the data collection was started, we intended for each sensor to record signals regularly every 10 milliseconds (i.e. a constant sampling frequency of 100Hz) based on a local timer on the sensor. We planned to then re-synchronize all data to common timestamps using the initial host clock start time and the stagger order of the sensors. After collecting the dataset, we determined that some motion sensors experienced cumulative errors in sample interval calculation. As a result, the motion signals were sampled at irregular intervals resulting in a variable sampling frequency across sensors (98 Hz on average). Despite the irregular sampling intervals, we had accurate timestamps resolved to 10 microseconds. After adjusting times using the stagger order, we interpolated the non-uniformly sampled data to a common 10 microsecond clock, then decimated the signal by selecting new samples spaced exactly 10 millisecond apart on hundredth-second boundaries. Resulting signals maintained the integrity of the original waveforms while establishing identical sample times across motion sensors, allowing simultaneous analysis across sensors.
>
> The embedded computing unit (NI myRIO) that recorded flow signals from laboratory equipment also connected to the host application via USB, synchronizing

> flow data to the same clock as the motion sensors via LabView libraries. We found it necessary to correct a drift in the myRIO clock relative to reference clock on the host device. Since we had programmed the system to (a) start and stop recording the kinematic and flow data at the same time, and (b) use identical sampling rate, we were able to quantify the drift using the difference in the number of samples recorded over the trial duration. The magnitude of the drift was 2.363 milliseconds per second of reference clock time and this was very consistent across all trials (standard deviation 0.0277 milliseconds per second of reference clock). Using this precise estimate of the drift, we synchronized the flow and kinematic data. [27]

Breathing signals can range from 3 BPM to 60 BPM, or 0.05 Hz to 1 Hz; a non-causal, Butterworth bandpass filter was used to isolate breathing motions (4th order high pass 0.05 Hz, 6th order low pass 1 Hz). See Figure 3.5 and 3.6 for examples of the filtering process and resulting data. Zero-phase non-causal filtering eliminates issues with phase shifts, particularly important for accurate comparison of signals with different frequency components across different sensors. While real-time implementations are desirable, predictive monitoring applications would not need hard real time results like feedback systems for robotic surgery would; filtering could easily be delayed by a long enough window to allow for non-causal filtering in a clinical setting. Resulting signals primarily showed prominent breathing motion, with some variation due to physiological response or gross motion noise. The filtered signals effectively capture two groups of information: (1) rate, timing and synchronization, and (2) amplitude and dimension of motion. Returning to the sensor model, band filtering eliminates the static sensor-frame bias and heavily attenuates the high-frequency components of the white noise process; if the scale factors and other small distortions are taken to be negligible, as in [88], then the model reduces to the time-varying components of the reference signals. Thus, the accelerometer catches motion and the dynamic component of gravity due to orientation change, the gyroscope catches the true angular rate, and the magnetometer catches the dynamic component of the local magnetic field. While they are all subject to in-band noise, the processing eliminates the need for static or dynamic calibration processes. This is non-insubstantial; calibrating seven IMUs would take much longer than the typical vital sign collection performed by a healthcare professional.

Due to motion effects of settling into regular breathing at the start and getting up near the end of each collection, filtered data were cropped before further analysis. Specifically, one-minute recordings starting 15 seconds after the start of the breathing interval were used for correlation and rate-accuracy analysis, as in [19].

3.3.2 Emergency Department/Hospital Ward Patient Data

Clinical validation was initially performed in the U.Va. Hospital Emergency Department, then expanded to include wards and critical care units throughout the hospital.

For ED trials, two-minute samples were collected every 30 minutes for between two and six hours, with a typical trial lasting four hours. An additional two-minute record was recorded at the end of the trial. The first and last samples were supervised by a research coordinator. Trials use the wireless apparatus described above, with the Android device mounted to the patient's bed rail. All patients had either a positive COVID-19 test or symptoms of respiratory distress when admitted. Sensors were left on the patient for the duration of the trial, with data downloaded at the end of the trial.

For trials on previously admitted patients, many of the same procedures were carried over, with the following exceptions: (1) all segments were supervised, (2) sensors were applied and removed for each collection, (3) segments were collected at less-precise intervals, typically 30-90 minutes apart, and (4) all patients were hospitalized with a severe marker of respiratory distress.

The same filtering procedure from the EPCL data was applied to the clinical data. One notable difference is the prevalence of noise from gross body motions. The supervision of the EPCL trials ensured that patients were staying still and sensors were attached whenever

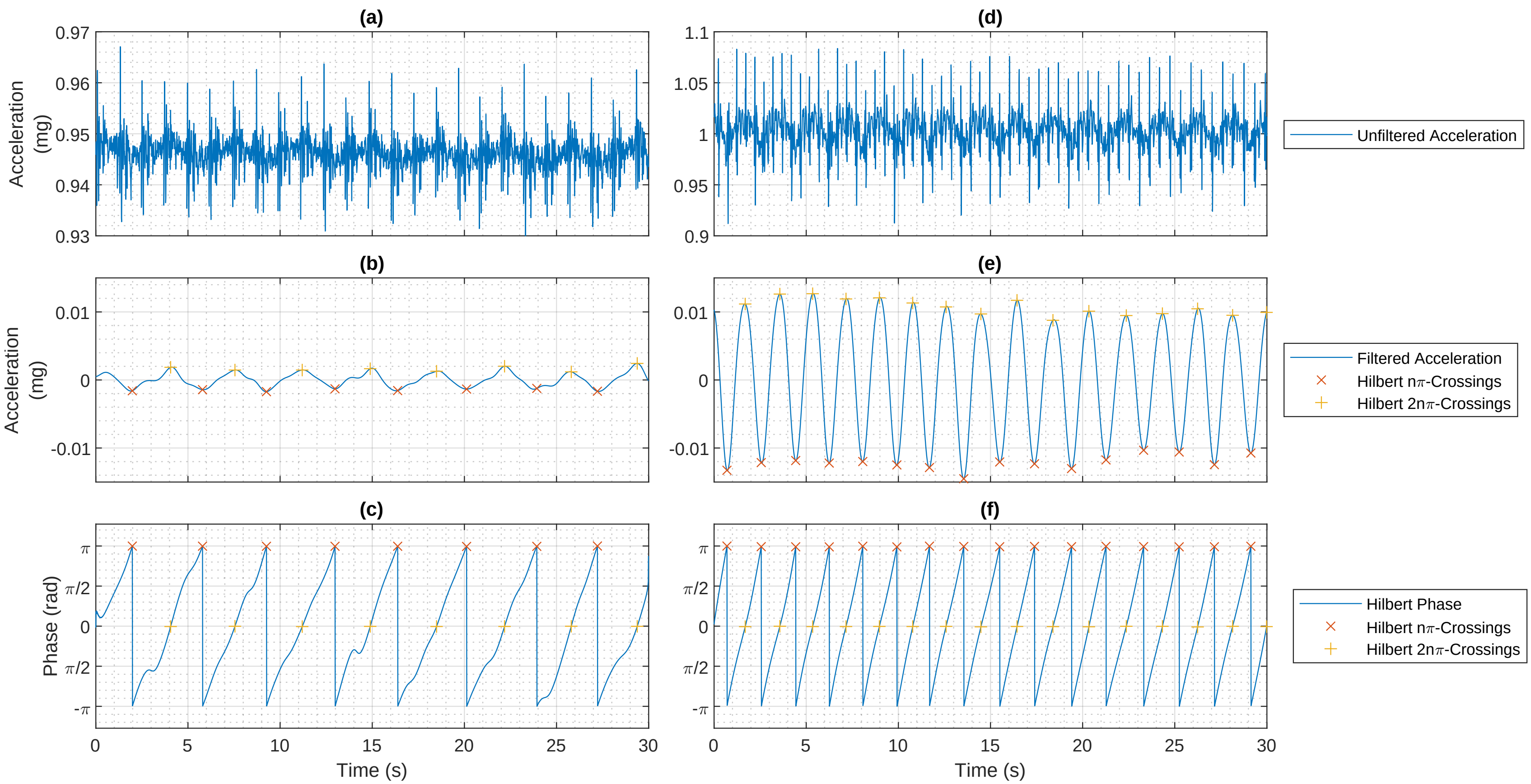


Figure 3.5: Filtering process for z-axis accelerometer for L_2 sensor at rest (left) and after max exertion interval (right), showing unfiltered acceleration with gravity ((a) and (d)), static bias, and noise, band-pass filtered waveform with Hilbert cardinal points ((b) and (e)), and corresponding Hilbert phase waveform of the motion of the chest at the 2nd rib ((c) and (f)).

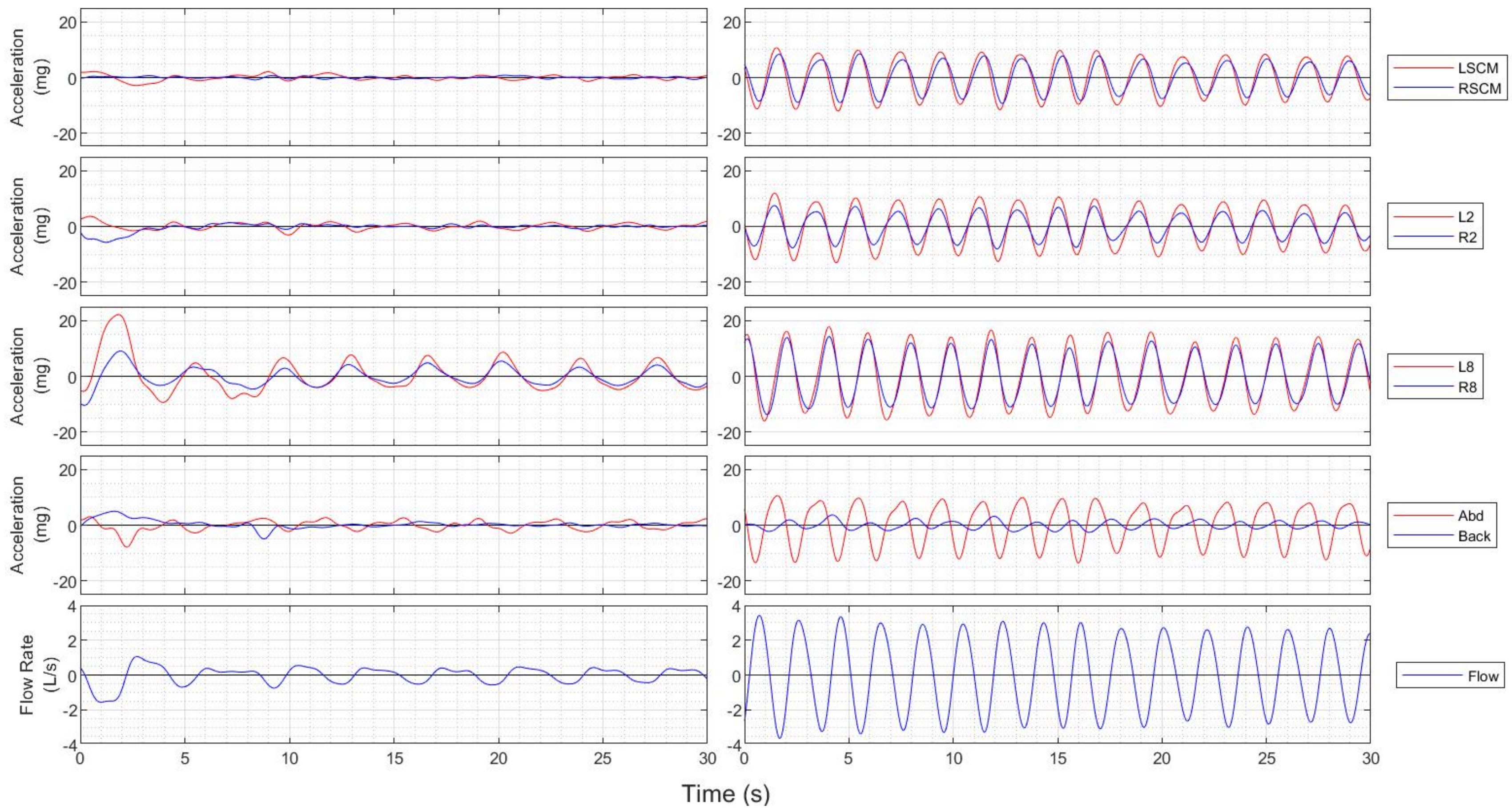


Figure 3.6: Example kinematic and flow data for healthy subject steady breathing (left) at rest (t = 0 to 30s after deep breaths) and (right) after peak exertion (t = 30 to 60s after deep breaths). As expected, respiratory rate increases (tachypnea) between rest and peak exertion with similar increases in the amplitude of measured accelerations.

possible. The semi-supervised paradigm in the ED means that resulting motion signals are substantially noisier, even with automated verbal reminders from the Android device. Noisier recordings necessited reliable methods; more data were included per collection with the understanding that gross motion noise would be detected and removed during analysis. To reduce startup artifacts while preserving the maximum data prior to additional segment filtering, final collections were cropped by 10 seconds on either end, resulting in recordings that were approximately 100 seconds long.

For the purposes of this study, outcomes were divided into three categories: (1) patients sent home from the ED, (2) patients admitted to a hospital ward for acute care, and (3) those admitted to an ICU. Patients who were admitted to a hospital ward before eventual transfer to an ICU were classified as ICU, and the single patient death is counted was also counted as ICU.

Chapter 4

Monitoring Respiratory Rate with ARK

The fundamental metric of respiratory kinematics is the respiratory rate as measured by the ventilation motions of the chest wall. While quantitative clinical application revolves around average rates, the heart rate variability research examined below provides motivation for similar work with breath rate variability. A review of current methods for finding the rate of a noisy, chaotic oscillator is given, along with constraints for a robust solution. The Hilbert Transform and analytic representation are introduced and their phase-tracking qualities described. To complete the rate analysis, the Berger algorithm is presented and used as a smoothing resampling algorithm, producing an instantaneous respiratory rate time series. The technique is experimentally confirmed; rates produced from limited inputs correlate reasonably well with instantaneous rates from flow data, leaving room for more intelligent approaches based on sensor fusion.

4.1 Problem Background

The problem of respiratory rate variability is an understudied topic due to the lack of established methods for recording and analyzing relevant signals with necessary accuracy and precision. Still, analogues exist within medicine that demonstrate the under-utilized potential of existing monitoring approaches. Heart rate variability examines the corresponding problem for heart rate sequences, attempting to learn more information about the dynamic process of cardiac impulse generation than a typical average pulse rate would yield [79]. With the advent of ECGs and particularly digital recording of large ECG databases, heart rate variability has seen significant research.

Typical shapes were described from the plethora of data, notably the sinus rhythm from the heart's sinoatrial node. Researchers found consistent landmarks of a sinus rhythm; the spike of the heartbeat signal, for instance, is denoted as the "R wave" of the "QRS" complex. The spikes are easily detectable, rendering intervals between R waves (R-R intervals) consistent, robust analytical mechanisms for analyzing changes in pulse rate on a per-beat resolution.

Resulting methods have examined ECG and other derived metrics for signs of heart rate variability. Measures developed include time-domain metrics like standard deviation of R-R intervals [107] as well as frequency domain analysis of heart rate series using an FFT [97] or autoregressive model [93] [92]. Graphical representations of sequence variability are achieved with Poincaré plots; each heart beat interval length is plotted against the previous, with highly varying sequences scattering over a larger area [108] [109] [110]. Finally, chaos-based quanitifcations of oscillator complexity with fractal mathematics [111] [112] and entropy analysis of successive RR-intervals [101]. These measures have been used to quantify risk of fetal stress [113], deterioration after a myocardial infarction event [114], neonatal sepsis [102], and more. Additionally, significant work has gone into understanding normal causes of heart rate variability; this includes the resting sinus arrhythmia [79], the interaction between the oscillation of respiratory system and correlating fluctuations in heart rate. The goal of these efforts and many more like them has been to apply useful technology to learn more about the underlying physiology. Effective monitoring effects better care by offering a more complete view of a patient, improving patient outcomes.

Corresponding problems in respiratory sensing have not been solved; current methods either do not adequately measure respiratory rate in a clinical setting or have not been adopted due to practicalities. Manual counts are common, prone to large amounts of error and with no ability to understand short-term variation [11] [12]. ECG- and PPG-based methods have been demonstrated to recover accurate respiratory activity, but neither offers any other functionality beyond their original and rate-derived metrics. An ideal respiratory monitoring system would be able to measure more than just rate, capturing the additional kinematic information that has been shown to be predictive without clear methods of measurement [18]. The historical perspective on heart rate variability suggests that such a system would have many clinical applications.

4.1.1 Methods for Extracting Rate

The ARK system is appropriately designed to measure respiratory rate. As a sensor moves with the chest, each axis of measurement oscillates with the rate of breathing. This amplitude-varying signal is isolated with the filtering described in Chapter 3. Current methods for extracting rate information from inertial sensors include spectral analysis, cycle counting via peaks or other landmarks, and combination time-frequency approaches.

Analyzing the frequency spectrum for the component with the highest power can yield an accurate measure of the rate when compared with the same metric from a baseline [73] [68]. Still, spectral analysis only captures a single estimate for a recording, with little room for understanding variation. While some analyses look at the shape of the whole spectrum [63], deriving a single respiratory rate estimate for any interval of time can be difficult. At best, spectrum-based techniques compress the information about the respiratory rate time series to a single average. At worst, the spectrum is overrun with noise and variation in the respiratory rate signal such that the average respiratory rate cannot be accurately estimated. Without delving further into the time domain, power frequency analysis does not contain the necessary depth to discern respiratory rate variability, and is limited in its ability to quantify the true rate.

Cycle counting methods quantify the average rate by counting cycles per unit time over some collection period[71] [67]. Cycles are defined by some landmark, most commonly the peak of a signal [115] [59]. Landmark detection relies on clean, sinusoidal-like signals or pre-processing beyond linear filtering [72] to provide meaningful output; more complex signals stump the often over-tuned or over-simplified filters. Common metrics like peak height, peak-to-peak distance, or peak prominence above local troughs require a cutoff parameter. Hard cutoffs struggle to find a good cutoff between sensitivity and specificity. Breathing is a highly dynamic process; the amplitude and frequency can change rapidly. To try to account for these variations with excessive decision rules or program-varied parameters raises the burden of proof that the resulting process is robust to noise or process variation. Peak detection methods have been used successfully in many applications, but the typical approaches do not lend themselves to theoretical robustness. They rely on a cleanliness of signal, which can be disrupted even if pre-processing is used. A data-driven approach might yield more accurate results, but that would require more data than currently exists. Peak detection does enable cycle-to-cycle analysis, which offers more room for examining variability than spectral techniques. Still, a more robust system for determining peaks or similar recurrent landmarks would be desirable to address the shortcomings of traditional peak detection approaches.

A mix of the above approaches would be peak finding in a time-frequency spectrogram [75]. Ideal candidates include the STFT [19] [116], CWT, synchrosqueezing transform [74] (built off STFT and CWT), SPWVD [55], and the Hilbert-Huang transform [76]. While this would allow for simultaneous analysis of the frequency components as they change in time, the limitation of entangled time and frequency resolutions would ultimately prevent accurate analyses in one or both domains. These techniques are also more useful for situations like musical spectral analysis, where multiple frequencies of interest and even regions of interest are expected and desired. The problem of respiratory rate is unique in that only one frequency exists at a given point in time, meaning that the simultaneous tracking of multiple frequencies is only useful for noise rejection. While the SPWVD method, synchrosqueezing transform, and Hilbert-Huang transform take different approaches to allowing the extraction of the instantaneous rate of multiple non-stationary components, they all suffer from an overabundance of information; at some point a single respiratory frequency component for each time index must be chosen, and only the frequency of that particular component matters.

An interpretation of the respiratory signals as phasors with a single time-varying frequency (i.e. only one frequency at each time) would be a more appropriate treatment of the data. Given that the signal band of interest is known (3-60 bpm), filters can be constructed to isolate the signal of interest. While harmonics are possible within this range, we can make the assumption that harmonic content will be dwarfed by steady breathing signals, and further that noise in the process can be processed by appropriate smoothing. The key basis for the following method of extraction of instantaneous rate is the use of known limits on the possible breathing signals (band-limited, single frequency) in order to produce a more reliable process with the same expectations of accuracy. Within this paradigm, more care can be given to rejecting noise in an effective and robust manner.

4.2 The Hilbert Transform and Analytic Representation

Analytic representation is a paradigm for tracking the progression of an oscillating signal by creating a phasor-like function with time-varying frequency and amplitude [27] [75]. The process converts a real signal with positive and negative frequencies to the corresponding complex analytic signal with only positive frequency components. The resulting progression in the complex plane allows for separate analysis of the time-varying amplitude and phase progression of a signal. Consider a real-valued signal $s(t)$. The analytic representation s_a of real signal s is defined as

$$s_a(t) = s(t) + jH(s)(t) \tag{4.1}$$

where j is the imaginary constant and $H(s)$ is the Hilbert transform of s. The Hilbert transform shifts the phase of all positive frequency components of s by $-\frac{\pi}{2}$ radians and all negative components by $\frac{\pi}{2}$ radians [117]. Multiplying by j phases all components by an additional $\frac{\pi}{2}$ radians, creating a complex pair with the original signal, with the sum having only positive frequencies as negative parts cancel out. The transform is calculated by convolving the original signal with the Cauchy Kernel, $\frac{1}{\pi t}$ [98]:

$$H(s)(t) = \fint_{-\infty}^{\infty} \frac{s(\tau)}{\pi(t-\tau)}\, d\tau = \fint_{-\infty}^{\infty} \frac{s(t-\tau)}{\pi\tau}\, d\tau \tag{4.2}$$

The continuous formulation of the Cauchy Kernel has a discontinuity at $t = 0$, so the principal-value integrals are computed by taking limits around the discontinuity. In practice, the Hilbert Transform is often computed discretely, allowing for numerical methods in place of analytical solutions. The Matlab implementation of the *hilbert* function, for instance, finds the two-sided FFT, zeros out negative frequencies, then computes the inverse FFT of the positive-only spectrum, as in [118].

With this in hand, the complex analytic signal can be reexpressed as a phasor (time argument dropped from all functions for clarity):

$$\begin{aligned} s_a &= s + jH(s) \\ &= \sqrt{s^2 + H(s)^2}\angle \tan^{-1}(H(s), s) \\ &:= A * (\cos\phi + j\sin\phi) \end{aligned} \tag{4.3}$$

Here A and ϕ represent the amplitude and phase functions of the complex analytic signal. For an unbiased signal, the amplitude envelope provided by A is more robust than peak detection and interpolation or integration. Both methods require some degree of tuning based on typical expectation of the input. Peak detection as a method is either noisy by over-detecting peaks or requires a minimum spacing between true peaks linked to the underlying expected rate. Methods using moving integration of RMS amplitude or similar require a time constant of integration fit to the underlying rate of oscillation. If a time constant is too small, the output envelope will be noisy due to artificial dips in between peaks of the original signal. If a time constant is too large, fast changes in amplitude will be lost in an overly-smoothed output. The Hilbert transform works without tuning for finding the envelope of an oscillating, narrow-band signal with widely varying frequency.

The phase function, ϕ, is particularly useful for tracking features of the periodicity and repetition of the typical signal shape. Consider the signal $s(t) = \cos \omega t$ for some constant frequency ω. In the frequency domain, the Hilbert transform of a single frequency shifts the original signal by $\frac{\pi}{2}$:

$$H(s)(t) = \cos(\omega t + \frac{\pi}{2}) = \sin \omega t \tag{4.4}$$

The resulting analytic representation of s is the sum of sinusoids:

$$s_a(t) = \cos \omega t + j \sin \omega t = e^{jwt} \tag{4.5}$$

which forms a circle in the complex plane. The landmark points of the original cosine wave (extrema and zero crossings) occur evenly at angles incrementing by $\frac{\pi}{2}$: the wave starts at a peak ($\phi = 0$), progresses to a downward zero-crossing ($\phi = \frac{\pi}{2}$), then a trough ($\phi = \pi$), then an upward zero-crossing ($\phi = \frac{3\pi}{2}$). The process then repeats on the following cycle from a wrapped angle of $\phi = 0$.

While these points could easily be found on a clean signal, detecting landmark points for non-sinusoidal, quasi-periodic waves becomes more challenging. Varying the amplitude of the wave disrupts peak detection methods looking for a certain amplitude threshold to start counting peaks. Avoiding false smaller peaks becomes difficult for highly volatile signals as the peak of the signal dips below a set threshold. Additionally, simultaneously varying

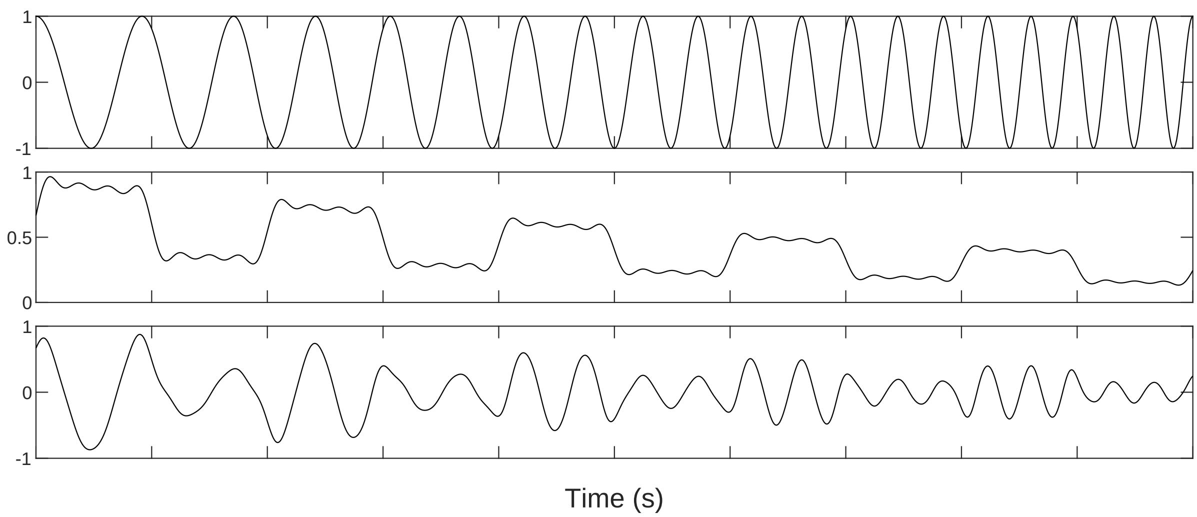


Figure 4.1: Composition of a signal with known time-varying frequency and amplitude. A chirp signal (top) with increasing frequency is given the decaying, oscillating envelope (middle), producing the resulting signal with time-varying amplitude and frequency (bottom).

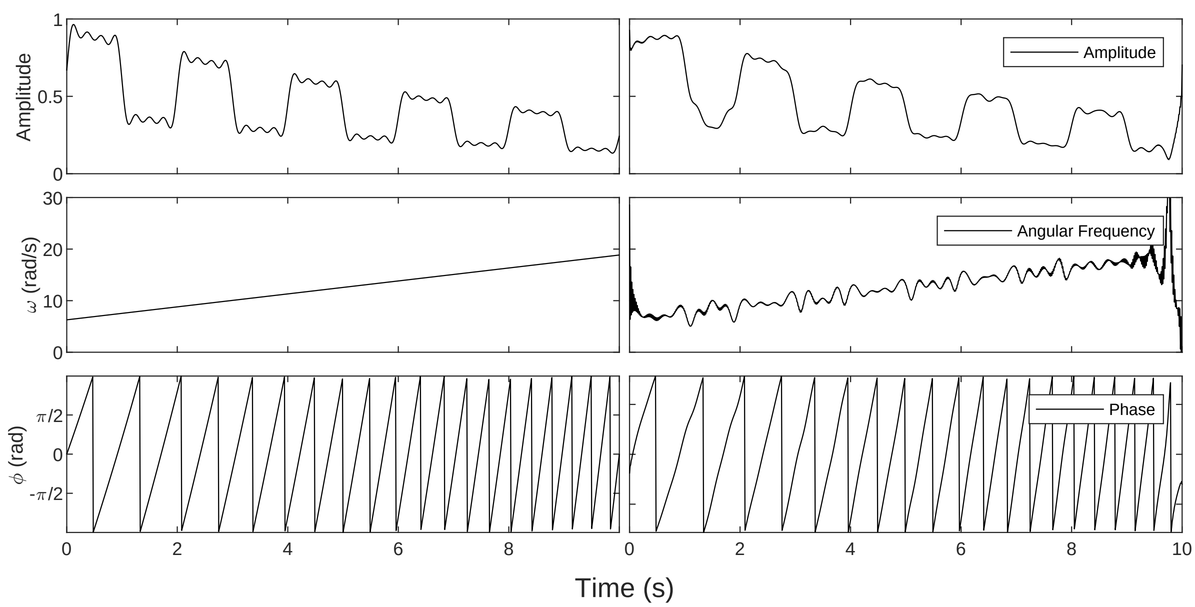


Figure 4.2: Example of the Hilbert Transform detecting dynamic amplitude and phase functions in the above signal. (left) Ideal amplitude envelope, instantaneous angular rate, and phase of the demonstration signal. The chirp signal has increasing frequency, starting from 2π and ending at 6π. The phase of the wave is computed as the integral of the frequency. (right) Approximations using the Hilbert transform and analytic representation; amplitude and phase are derived directly, while the instantaneous angular frequency is calculated with the first differences of the phase function. While there is some noise and lack of precise shapes, the curves clearly track the reference waveforms.

the frequency of the wave layers on concerns about detecting peaks that are separate cycles but unusually close in time. Searching for peaks with a minimum expected distance between peaks limits the frequency that can be detected. This distance could be optimized by roughly computing the frequency before looking for extrema, but such a step is unnecessarily complicated and prone to error.

The Hilbert phase function provides a computable alternative for finding quasi-periodic landmarks, extending the common notion of sinusoidal phase to narrow-band, quasi-periodic signals. Consider Figure 4.3. The phase function ϕ perfectly tracks the progression of the cycles, despite changes in frequency and amplitude. Similarly to the cosine wave, landmark points can be extracted at phases incrementing by $\frac{\pi}{2}$. By definition, the real-valued signal is a projection of the analytic representation to the real axis. Wrapped phases of $\phi = 0$ and $\phi = \pi$ on this projected axis track consistent points near the amplitude extremes on the original signal. Phases of $\phi = \pm\frac{\pi}{2}$ lie on the imaginary axis; these points fall where the signal changes sign. Thus, the Hilbert transform consistently identifies landmark points on dynamic signals, making it suitable for tracking the progression of respiratory signals (see Figures 4.1 and 4.2).

4.2.1 Comparison of Average Breathing Rate Methods

The Hilbert transform provides consistent landmarks at extrema, which can be used to compute the average rate over time, much like peak detection. Figures 4.4, 4.5, and 4.6 provide a comparison between three peak detection methods (minimum amplitude, minimum inter-peak distance, and minimum peak prominence) and the landmarks from the analytic representation. Notably, the analytic representation performs with small errors on each result, but the other filters stumble considerably. Both the minimum peak method and minimum peak prominence flounder with unstable peaking or low amplitudes. The minimum peak distance has similar troubles, with highly-varying rates and multiple peaks per cycle complicating the selection of a single threshold. Other popular algorithms (e.g. typical

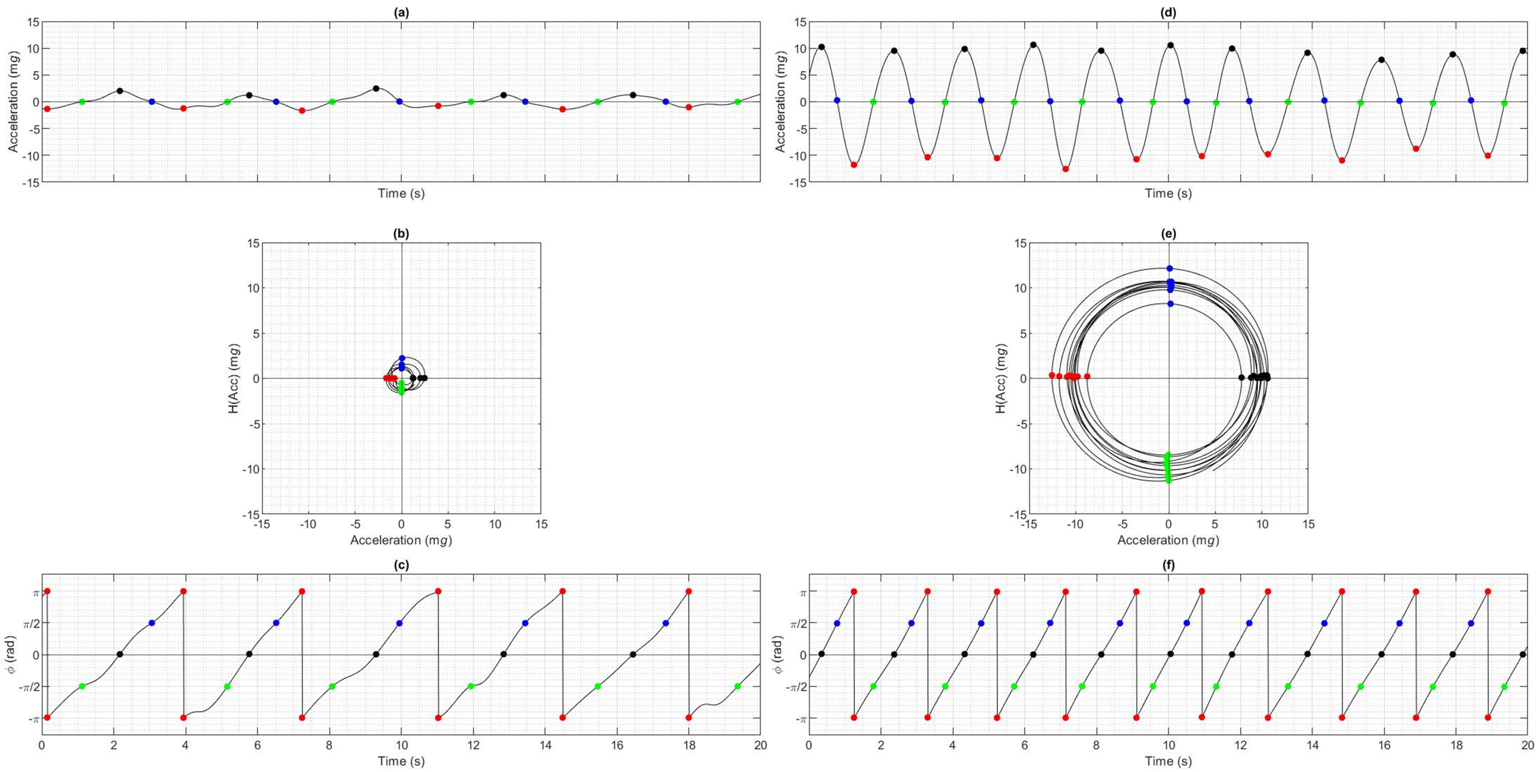


Figure 4.3: Example of the landmark identification process on sensor data, shown at rest on the left and after exertion on the right. (a) and (d) show the z-axis acceleration in the time domain. (b) and (e) show the analytic representation of each real signal, with the Hilbert transform plotted on the imaginary axis. (c) and (f) show the phase function, tracking the progression. Landmarks at $\phi = 0$, $\pi/2$, π, and $-\pi/2$ identify the important points in each cycle of the quasiperiodic waves.

dispersion algorithms) require more complicated parameter choice processes, all built to typical forms of signals. Further assumptions about the data would allow for custom-designed filters, but doing so could jeopardize robustness and reliability of the resulting method in the face of wide variation of wave shapes from breathing motions.

Of the peak detecting options presented, the analytic representation is the most robust method for tracking of the landmarks of the breathing process. The landmarks found are not always true peaks, but they mathematically correspond more to the point of repetition by definition of the analytic representation, and are very close in practice. Due to the separation of amplitude and phase functions, the progression can be measured even under low amplitude conditions. Additionally, the processing handles varying rates and amplitudes extremely well as a natural extension of time-varying phasor model. By inspection, the phase function pi-crossings were chosen as the landmark of interest, as troughs tended to be more consistent by physicians eye.

4.3 Instantaneous Respiratory Rate Resampling

Consistent landmarks can be converted into the desired instantaneous rate by applying a resampling process to the unevenly sampled breath interval information. Landmarks track cycles of the wave as phase progressions of 2π. The time between two landmarks becomes an estimate of the period over that interval, understanding that the dynamic rate process is changing over time. Computing the breath intervals for a continuous segment yields a step function of the breathing rate, transitioning between rates at each landmark timepoint. A more desirable series would be uniformly sampled, and appropriately smoothed.

Berger et al. [119] presents a method for converting interval series into a uniformly-sampled rate series. The method was originally designed for finding an instantaneous heart series with optimal frequency-domain representation, but is just as applicable to the landmark intervals identified here on breathing signals. The algorithm is repeated here for con-

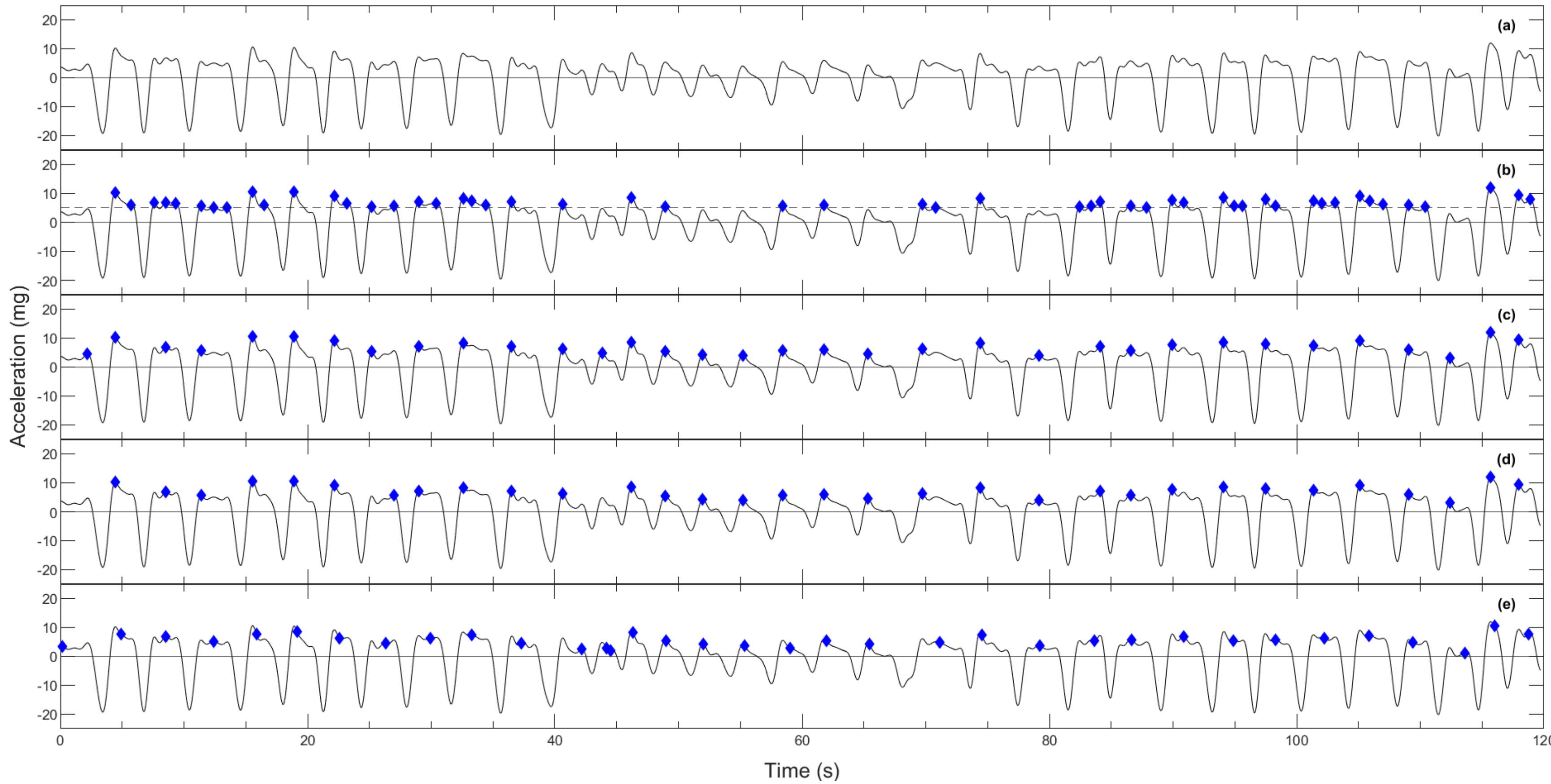


Figure 4.4: Comparison of peak detection methods for a signal with varying amplitude. (a) shows the raw signal. (b) through (e) are peak detection methods: (b) hard threshold at 5mg, (c) minimum peak distance of 2.2 sec, (d) minimum peak prominence of 10mg, and (e) zero-phase points from the analytic representation. Here, (c) provides the best fit; no height threshold can be chosen to catch short peaks and reject tall double peaks. (d) misses a single short cycle, while (e) has only one double-count. [27]

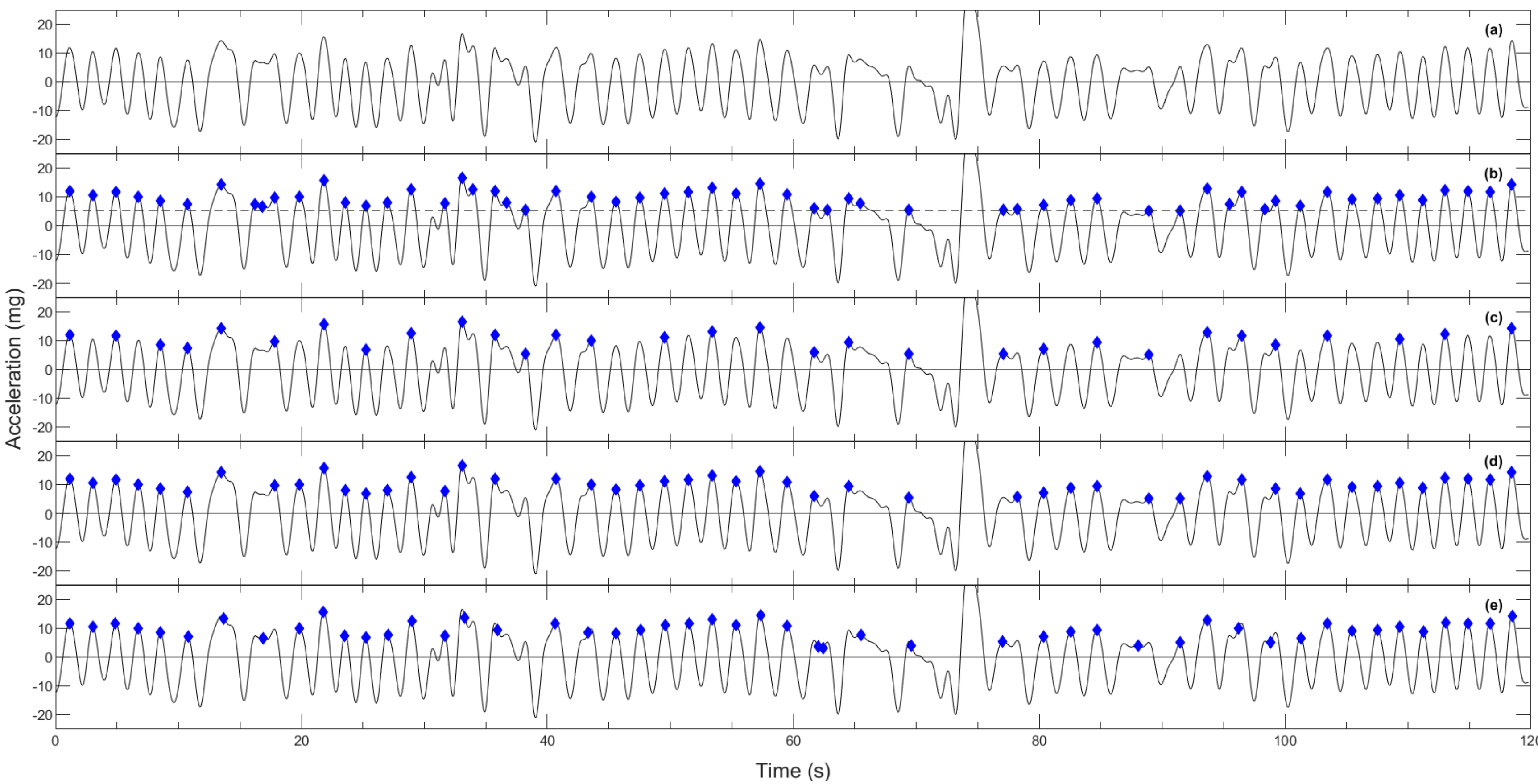


Figure 4.5: Comparison of peak detection methods for a rate-unstable signal ((b) amplitude threshold, (c) peak distance threshold, (d) prominence threshold, (e) analytic representation). (d) works the best here, with prominent peaks. (b) has multiple peak issues, while (c) is tripped up by the widely varying rate, with both a double peak and missed peaks. (e) again has only one double peak.

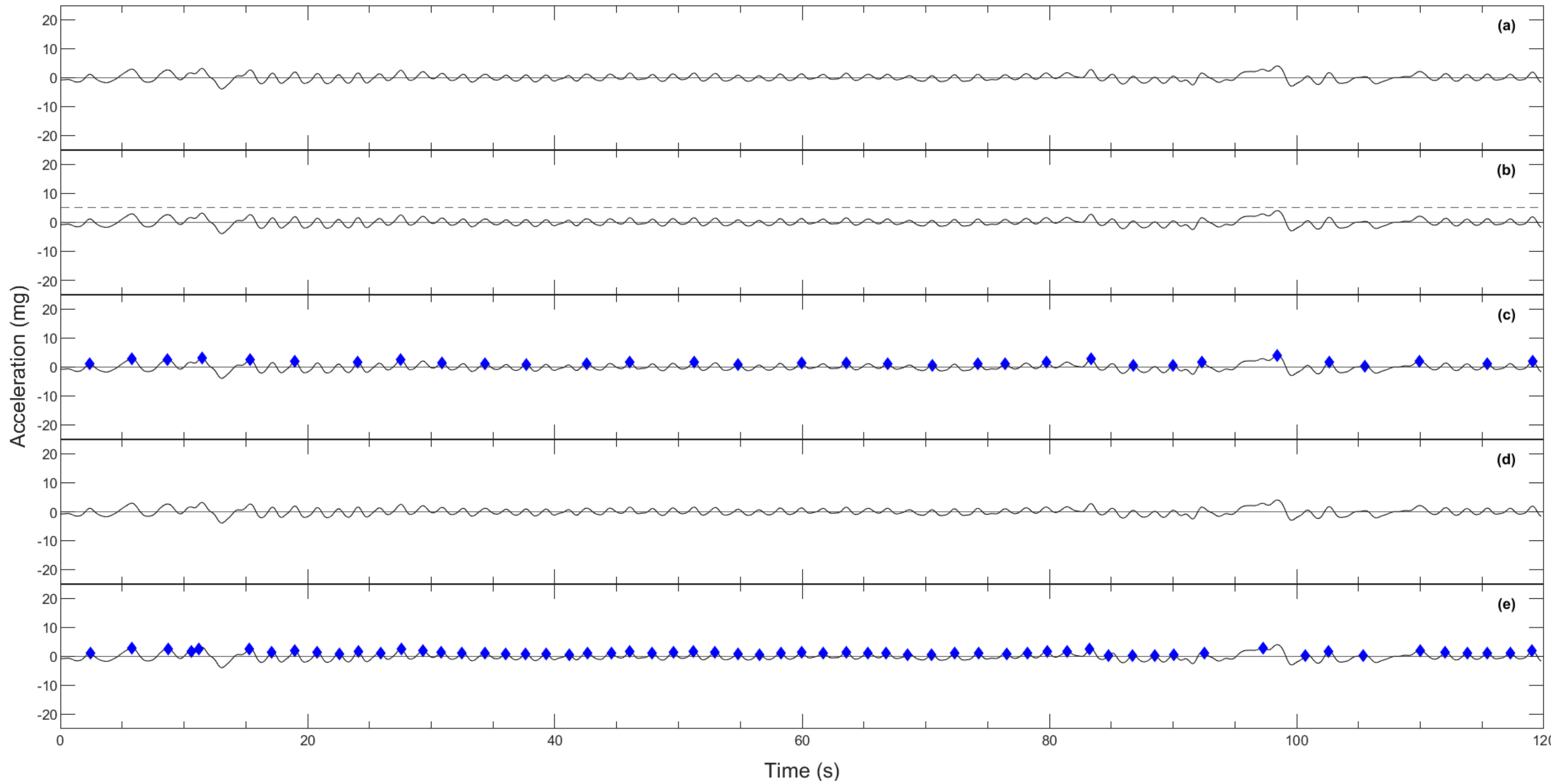


Figure 4.6: Comparison of peak detection methods for a low amplitude but steadily oscillating signal ((b) amplitude threshold, (c) peak distance threshold, (d) prominence threshold, (e) analytic representation). Of the standard peak detection methods, only the minimum peak distance filter registers any peaks, missing approximately half. The amplitude of the signal is too low for thresholds that have already missed cycles in Figures 4.4 and 4.5. (e) tracks the signal almost perfectly, with only one double cycle.

venience: (1) Landmark points are identified on a well-sampled signal; (2) A new timescale with uniform spacing between points is chosen; (3) A window is constructed for each new timepoint, spanning from the previous sample time to the following sample time; (4) The number of intervals in each window are counted, including fractional intervals; (5) The rate r_t at time t is calculated as

$$r_t = f_r * \frac{n_t}{2} \tag{4.6}$$

where f_r is the newly chosen sample rate for the resampled series, and n_t is the number of intervals in the window centered at time t.

This method performs a weighted averaging operation on the rates across the sampling window, as shown in Figure 4.7. This feature is ideal for smoothing the transitions in the interval time step wave. After reviewing various sample rates, we selected 1Hz as a balance between quickly changing rate information and appropriate smoothing between steps in the interval series. Higher rates did not produce desirable transitions, while lower rates missed meaningful variation in the resampled rate series.

The above Hilbert-Berger method can be viewed as a simplification of the Hilbert-Huang Transform and the Synchrosqueezing Transform. Rather than computing multiple mode functions and isolating the function of interest, the Hilbert-Berger process assumes a single mode that is already dominant, such that frequency representations of different instants will only have one frequency component that is consistent over time. The Hilbert-Huang Transform and Synchrosqueezing Transform both produce smooth frequencies by isolating individual components that are likely to be independent; the noise of tracking the instant frequency is reduced because the functions were specifically derived to have a well-defined progression of different frequencies. Here, the smoothing process is achieved by a windowing function for the cycle-counting component of the Berger method. While this still can result in noisy estimates, there is a limit to the effect on the output due to the averaging; as seen shortly, the process can work quite effectively even under dynamic and noisy conditions. In terms of the robustness, the lack of over-complication in the method ensures that it will

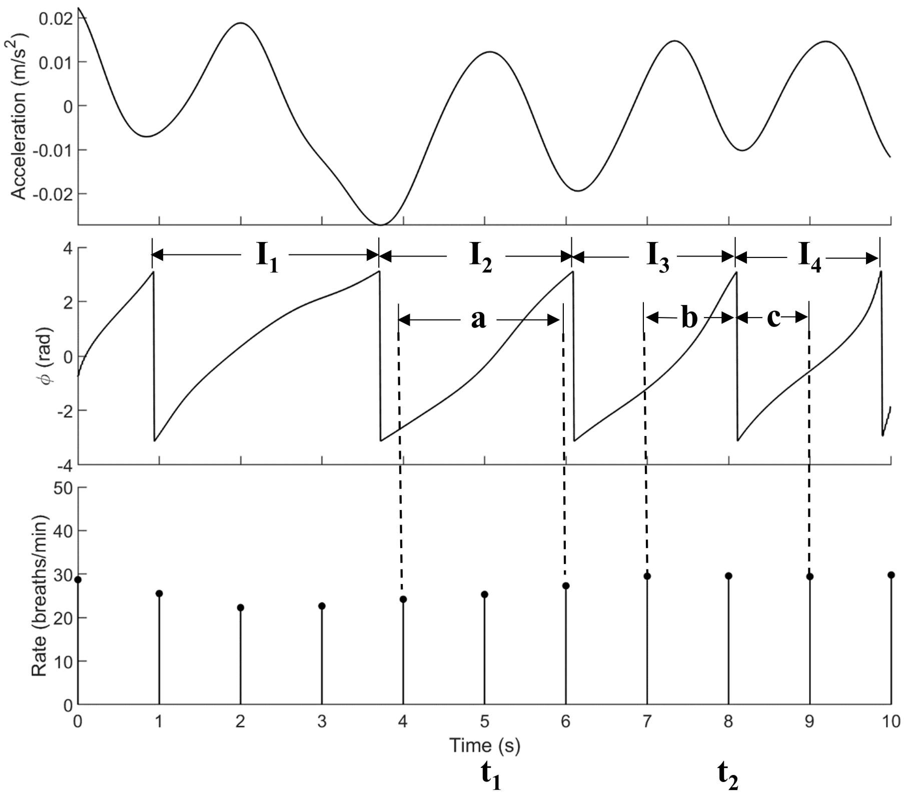


Figure 4.7: Breakdown of the respiratory rate resampling process. The top two curves show a segment of the acceleration signal and its instantaneous phase. The respiratory rate samples derived from these intervals are shown at the bottom. First, breath intervals (labelled as I_1 to I_4) are determined using consecutive occurrences of a phase angle (π radians, in this case). Next, a sampling rate f_r for the respiratory rate signal is chosen as desired (1 Hz in our case), without regard to mean respiratory rate or sampling frequency of the acceleration signal. For each sampling point, we count the number of breath intervals (n_i), including fractions, that occur in the time window extending from the previous sample to the next. For example, at time t_1, $n_{t_1} = \frac{a}{I_2}$ and at time t_2, $n_{t_2} = \frac{b}{I_3} + \frac{c}{I_4}$. The respiratory rate ($r_i$) at each sampling point is calculated as $r_i = \frac{f_r n_i}{2}$. [27]

work exceptionally well except when the underlying signal is too noisy to contain a single, meaningful rate. If a signal is that noisy to disturb the rate process, it would likely be a motion error that would confound other methods trying to find a single frequency in the noisy spectrum. A more sensible approach would be to determine an algorithm to identify the intervals of noise and remove them from computation.

This argument is based on Occam's Razor, namely that the method with the fewest assumptions should be tested and used first. An important corollary of the razor is that when methods are equivalent for reaching an objective, the simplest method is typically the best because the success and failure of the model are more easily described. This is important for a clinical application, where the data produced have the potential to both help treatment when used correctly and hinder treatment if the data are inaccurate or misused.

4.4 Evaluating Instantaneous Rate Signals

Figures 4.8, 4.9, 4.10 demonstrate the potential of the Hilbert-Berger measurement process. Time-series are well recorded by most sensors, rising and falling as appropriate. Some sensors do make mistakes, either spiking due to noise or over-fluctuating due to an interval measurement with jitter. Still, the ability to accurately track the rate is clear. Further, the sensors correlate well when compared with average rates from the flow signals. Most impressively, the mean operation across all three sensors improves the rate estimate across the entire range.

Signals were evaluated using cross-correlation, root mean square (RMS) error, and percent error against the instantaneous rate from the gold-standard flow rate signal, also derived via the Berger algorithm. Error statistics for the time-series analysis are summarized in Figure 4.11. Sensors were all somewhat noisy, between 14% and 18% error, while the mean series improved performance to 13.2%. Though the true rate would be unknown in a clinical setting without flow, the best individual sensor offers an interesting perspective and an

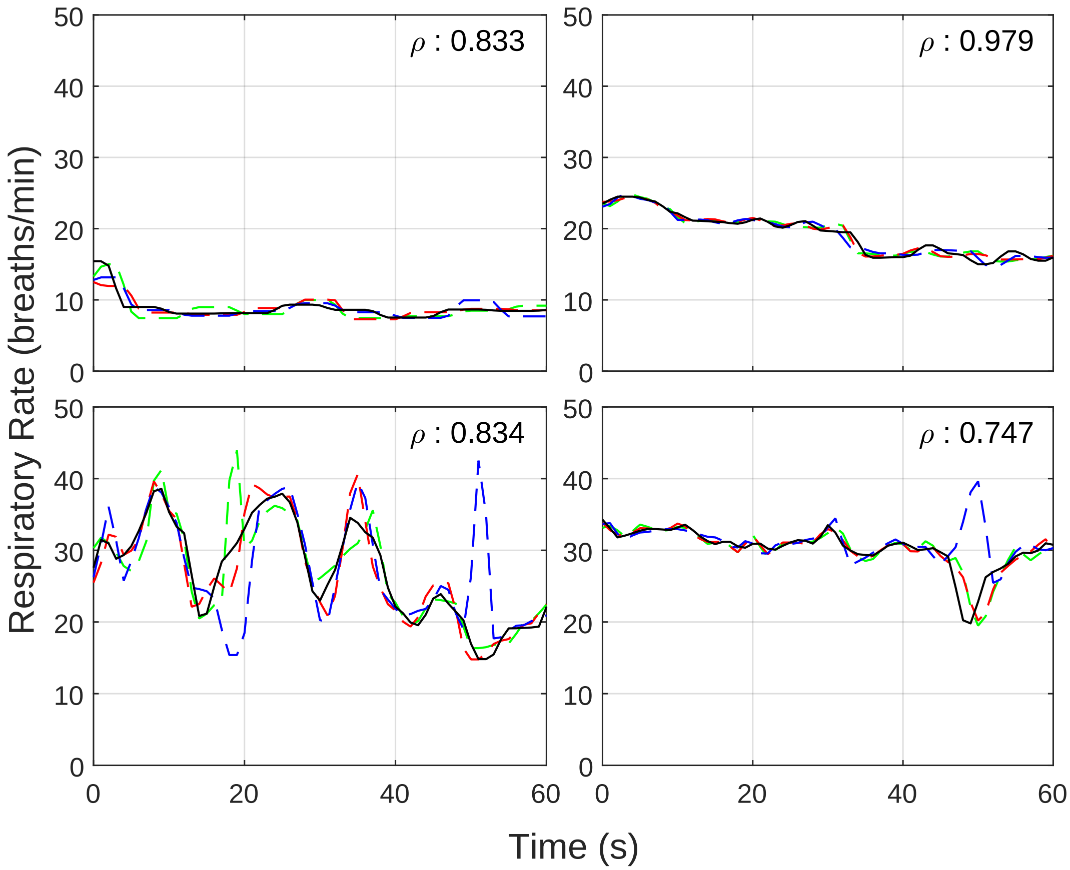


Figure 4.8: Comparison of four one-minute segments of data with a variety of recorded rates. The resampled curve from the gold-standard flow in black; dotted lines represent 8th rib (red and green) and abdomen (blue) signals. (a) shows a resting signal with a low, steady rate. Panels (b) through (d) show exhaustion: a smooth recovery in (b), cyclical fluctuation during recovery in (c), and an interruption in recovery in (d). The sensors follow the flow rate curve with high fidelity, resulting in strong correlations. [27]

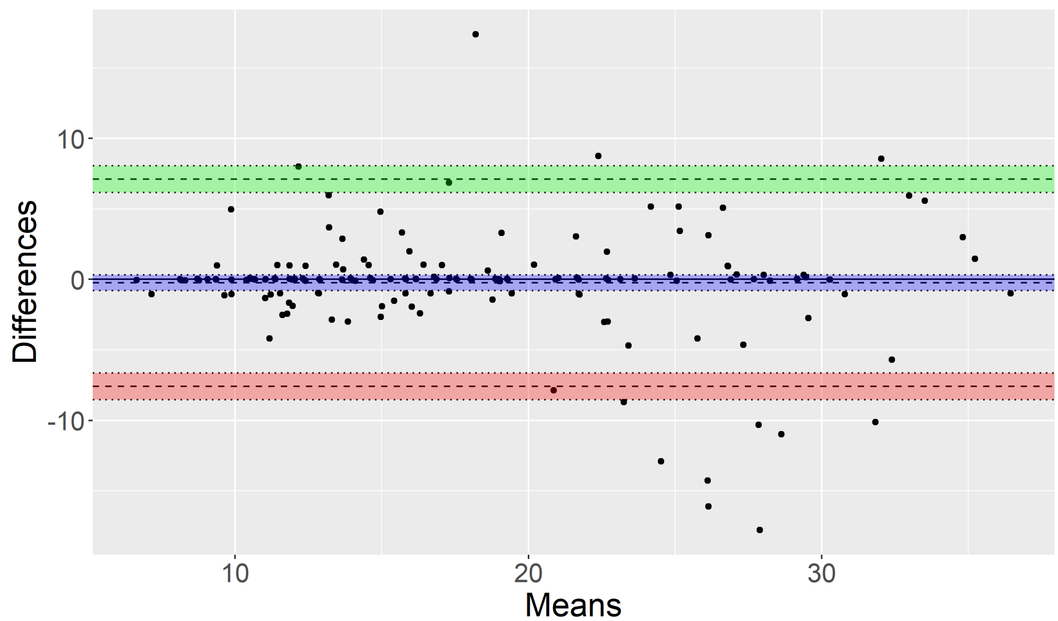


Figure 4.9: Bland Altman Analysis of select EPCL sensors showed a bias of 0.1 bpm and 95% Limits of Agreement of 7 bpm.

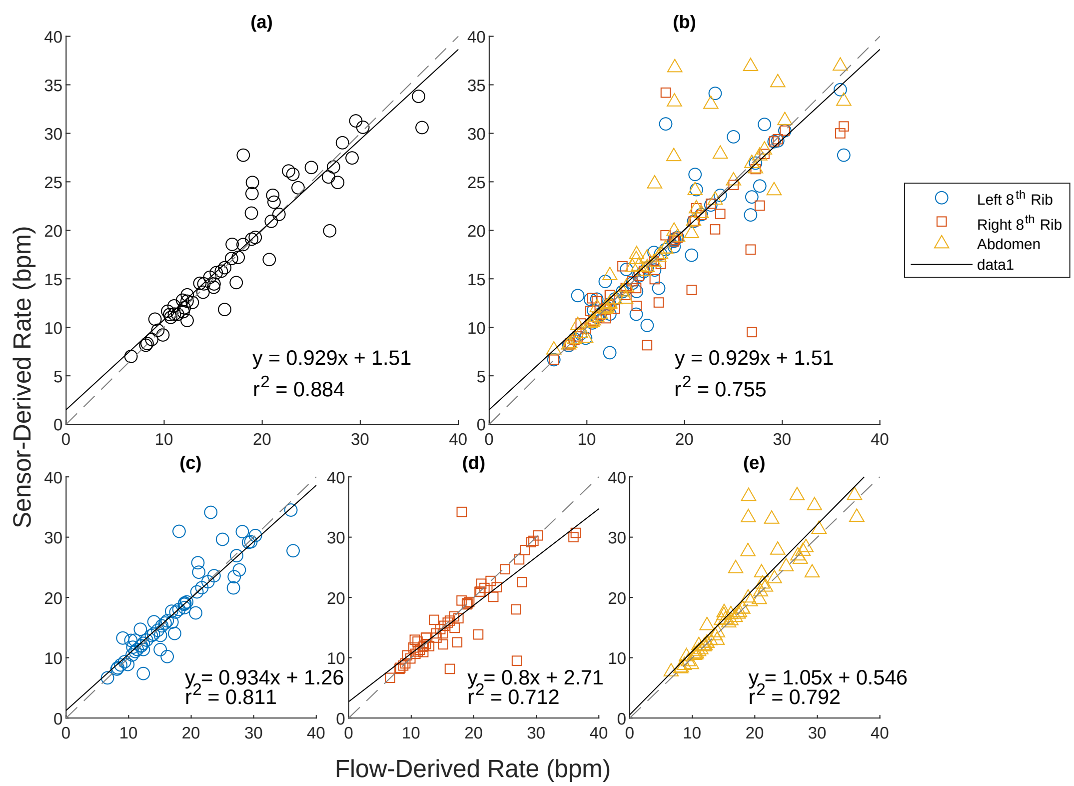


Figure 4.10: Correlation plots for sensor-derived flow rates: (a) shows the mean of recorded rates, while (b) shows all rates from individual sensors. (c) through (e) breaks down the collection by sensor type, showing (c) left 8th rib, (d) right 8th rib, and (e) abdomen.

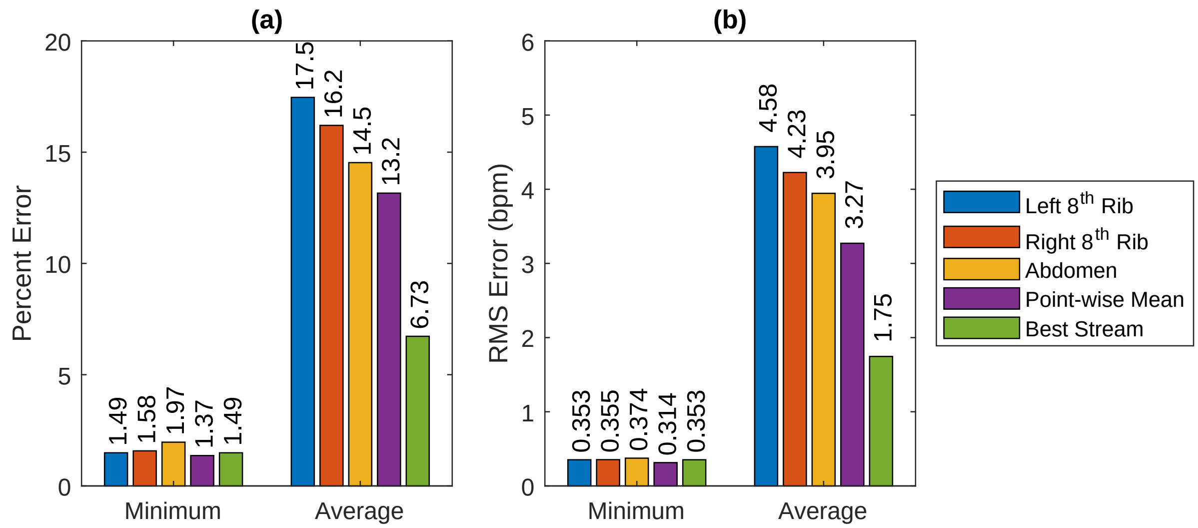


Figure 4.11: Comparison of time-series errors between sensor estimates and flow-derived values. (a) shows percent error metrics, while (b) shows the RMS error versions. Sensors had similar errors, while the mean time-series offered a slight improvement. With knowledge of the best individual signal per segment collected, error dropped to almost half. The time-series mean still had the best performance over all segments.

attractive performance mark: 6.73% error, nearly half the mean series mark. An algorithm to detect the best sensor by some criteria could be a useful tool for improving performance. That said, the best performance over all segments was the mean series signal. Combining sensor estimates on a per point basis has the potential to out-do any one sensor collection.

Sensors had high potential for correlation and accuracy, with cross-correlation coefficients as high as 0.99, mean relative errors at 1.49%, and RMSE of 0.302 bpm (measured on a two-minute segment of data). While single-sensor performance had a high ceiling, motion noise across the torso had a considerable effect on the quality of data. Data without motion noise can be cleanly tracked through the combination of the Hilbert transform and the Berger interpolation algorithm, but motion disrupts the filtered data stream to the point where Hilbert landmark points are inaccurate. The abundance of data on the ARK platform suggests that a larger approach that examined the relationship between sensors and their produced data streams could be capable of synthesizing a single rate signal for the entire breathing kinematics system with improved accuracy over the component signals.

Chapter 5

Similarity Filtering Across ARK Sensors for Robust, High-Fidelity Respiratory Rate

While a robust method for producing an instantaneous rate from a noisy signal is useful, it makes no use of the combined knowledge of using three multi-axis sensors on multiple locations. One could apply the process to all the streams and take an average; this would likely be accurate for reasonably clean signals, but the process would fail under sufficient noise and lose accuracy. Here, a method for deriving a single respiratory rate from multiple estimates is given, based on modeling of the differences between pairs of rate measurements. The method is evaluated against the instantaneous rate from the flow data and found to be highly effective at finding both average and instantaneous rates. Further, the variability of the sequence is quantified as the standard deviation of respiratory rates and correlation is similarly demonstrated, with good agreement between the sensor and flow measures.

5.1 Problem Background

The method in the previous chapter is effective for deriving a resampled rate series from clean signals with varying amplitude and frequency. Motion sensors like the inertial measurement units presently employed, however, are notoriously noisy [10], with gross motion from body movements often dwarfing the breathing signal of interest. Even slight movements can distort signals enough to break quasi-periodic trends. Tracking and filtering out these errors is crucial to the success of any algorithm that seeks to extract underlying breathing rate information from inertial motion sensors.

The ARK system records a plurality of motion signals from across the chest. An ideal filter would somehow combine rate estimates from each signal on each sensor to identify the gross motion effects and exclude noisy intervals from further analysis. An intuitive approach would look at the similarity between signals from the perspective of correlation; signals recording similar respiratory content would register as higher correlating. Unfortunately, this approach runs into two issues: (1) a correlation procedure that produces correlations between two signals at multiple times requires input parameters of a maximum lag and a signal window size, whose selection is confounded by signals with periods ranging from one second to 20 seconds, and (2) gross motion noise on a short enough interval can easily correlate with similar excessive noise from another sensor, reducing the effectiveness of any noise filter based on correlation. A better filter would have a well-reasoned parameter derivation and directly try to isolate useful information and reject noisy outcomes.

5.2 Similarity Across Sensors

The rate-resampling algorithm was applied to each stream of data from each sensor. For data from the the ED trials, six sensors (back excluded) provided three axes of accelerometer, gyroscope, and magnetometer each, resulting in 54 resampled signals. After reviewing visual records for each signal across trials, the magnetometer was determined to be too noisy to extract consistent rate information, leaving 36 streams. The resampled timescales were based upon the original, synchronized timings, meaning that all 36 resample rate series for a given two-minute segment were identically sampled. This allowed for direct comparison of the rates produced by each stream and sensor.

Due to startup effects from the filtering and resampling process, resampled signals were cropped by ten samples on each end. Resulting signal groups were approximately 100 seconds long.

At each timepoint in a given 100-second collection, individual rates were assumed to be either valid measurements of the rate process or noisy estimates disturbed by motion. Since all valid signals can be assumed to be locally congregated, the problem can be cast as an instance of 1-dimensional clustering. While it would be sensible to extract a cluster based on the grouping of the data, the method becomes overly complex and ill-fitting in practice. First, traditional clustering methods are built to identify multiple clusters, whereas there exists just one good rate cluster; clustering noise would have no meaning and potentially impact the results of finding the true rate. Second, any algorithm clustering has some specification on the number of detected clusters, the size of clusters, or a related parameter that controls the clustering process. Kernel density estimation methods, commonly used to estimate probability distributions, compute sums of kernel functions located at each data point. Clustering methods based on kernel density estimation look for maxima and minima in the estimated function, but the quality of the fit is controlled by a free-ranging smoothing parameter. Additionally, the resulting clusters may or may not be located in a usable way with a clear mean of good signals.

A better approach is to use problem-specific information to find a meaningful true rate signal out of the set of resampled streams. Valid rates occur close together; invalid rates will generally fall much farther apart. Knowing that valid rates will be grouped together, the probability that any two are valid if they are close high, since the probability that they would be close and one or both would be invalid would be low. Similarly, the probability that both are good if they are not close is low, as they should be centered around the same true rate. Lastly, the probability that one or more is a bad estimate given the two are not close is high, since the range of rates for bad estimates is much larger than the local neighborhood of the true rate. Thus, the collection of pairs of sensors can act as a distributed measure of the closeness of the signals.

Two problems remain for a pairwise closeness analysis: (1) the exact meaning of "closeness" must be defined, and (2) a federating approach must be designed to combine distributed estimates of closeness into a single cluster that describe the true rate mean.

5.2.1 Evaluating Closeness of Paired Rates

The ideal definition of "closeness" for this particular problem is entangled in the pairwise difference between any two signals. Each has some expectation of typical performance if measuring the true rate, and again some expectation if sufficiently noisy. Let $p^{ij}(x)$ denote the probability density function for the absolute pairwise difference $x = |r_i - r_j|$ between sensors i and j, $p_a^{ij}(x)$ be the probability of agreeing rates, and $p_d^{ij}(x)$ be the probability of disagreeing rates. Interpolating between the two to gives the full distribution:

$$p^{ij}(x) = \alpha_{ij} p_a^{ij}(x) + (1 - \alpha_{ij}) p_d^{ij} \tag{5.1}$$

where $\alpha_{ij} \in [0, 1]$ is the interpolating factor. Intuitively, the disagreeing distribution has a considerably larger tail, as virtually all large differences are caused by large errors. Define the closeness threshold of any two sensors to be the point of equal probabilities of error between the two parts. If an absolute difference is below the threshold, it is more probable that it is drawn from the distribution of both-valid rates; if above the threshold, it is more likely to that at least one sensor is noisy. Still, calculating such a cutoff would benefit from modeling each of the components.

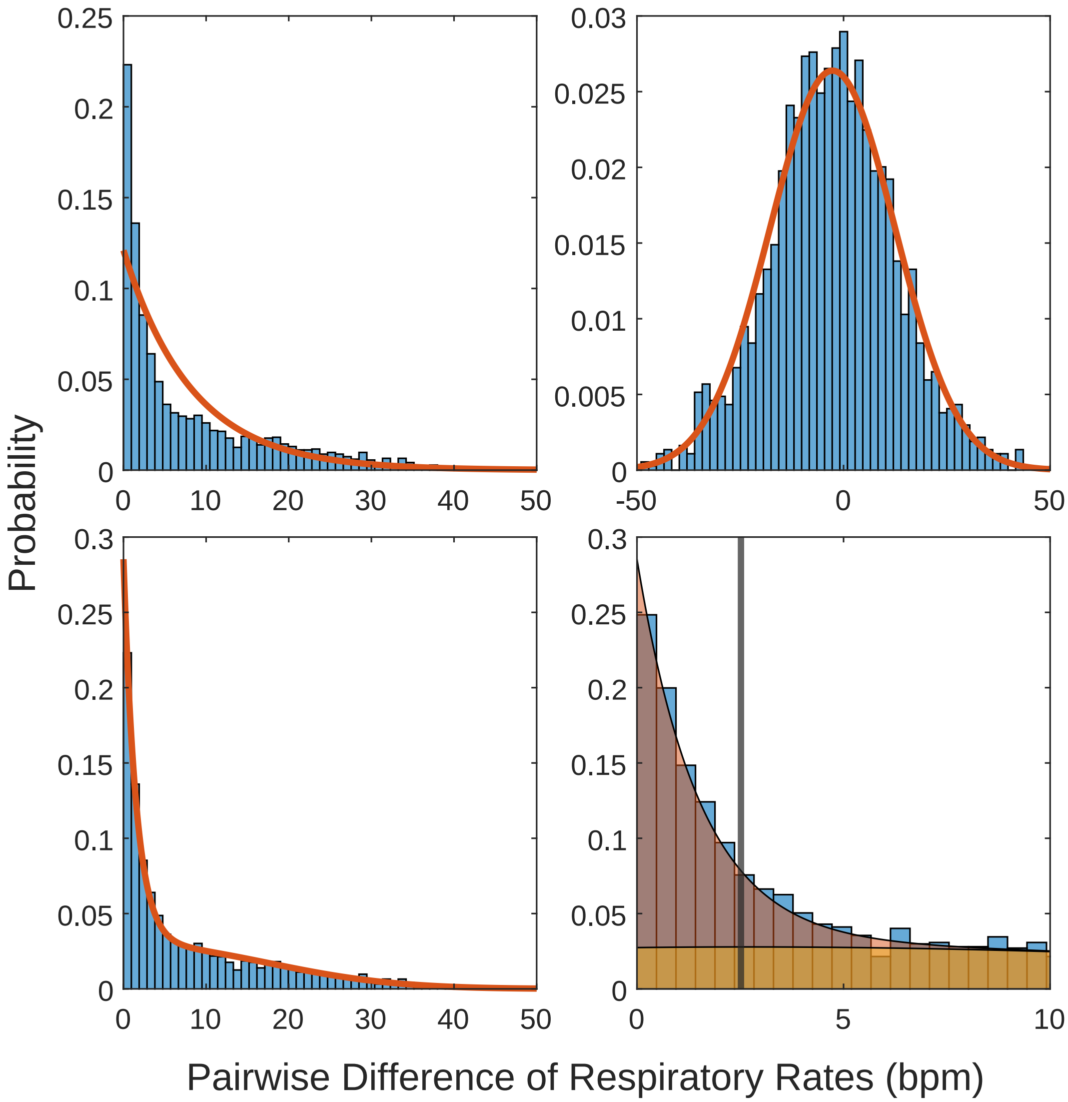


Figure 5.1: Process of fitting a probability density function to the pairwise difference of rates from a pair of sensors. (a) shows inadequate fitting of a single exponential to the absolute differences; (b) shows the distribution of unrelated pairwise differences, fitted with a normal curve; (c) shows the much closer fit of the combination of truncated normal and exponential; (d) shows the cutoff derived by balancing the weighted probabilities of error between the distributions, with component distributions highlighted.

Modeling these distributions exactly would require costly manual labelling of rates as either meaningful or not meaningful, dependent on the exact algorithm used. It would be prone to human error, and would need to be repeated for every change in the filtering and resampling process. To solve this issue, the distributions can be estimated based on existing data.

To model the overall distribution, samples of simultaneous pairwise differences between sensors i and j were collected and visualized. Each histogram of related pairwise absolute differences steeply drop off near $x = 0$. It is therefore assumed that agreeing rates follow an exponential distribution:

$$p_a^{ij}(x) = \begin{cases} \lambda_{ij} e^{-\lambda_{ij} x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{5.2}$$

where λ_{ij} is the exponential constant, equal to the reciprocal of the expected value of the distribution. A naive fitting of a single exponential ($\alpha_{ij} = 1$) to the whole distribution fails to capture the full shape of the curve (see Figure 5.1a). The exponential model drops off too quickly, leaving a wider tail that disrupts the fit. This confirms the need for a second distribution with the exponential one; the exact exponential constant and interpolating factor can be optimized given a shape of the disagreeing distribution.

To model the disagreeing distribution, the distribution of unrelated pairwise differences was visualized for each pair. Unrelated pairwise differences were calculated by sampling paired rates from all collections and trials, shuffling the samples, and computing the difference of shuffled rates. This is not an exact model for the distribution of disagreeing rates, as some rates used are not noise. Rates measured follow a loosely normal curve, meaning that pairwise differences tend towards the difference of averages. Still, it provides a decent proxy, and as shown shortly, works perfectly well as a model for the remaining distribution.

The resulting histogram appeared normal as well (see Figure 5.1b); to model the data, the maximum likelihood estimate of distribution parameters was estimated. Maximum likelihood estimation (MLE) is an apt tool for optimizing arbitrary distribution parameters, where likelihood over n samples is defined as the product of probabilities:

$$L(x, \theta) = \prod_{i=1}^{n} p(x; \theta) \tag{5.3}$$

where θ is a candidate set of distribution parameters. For a normal distribution, the MLE parameters are simply the mean and standard deviation of the data.

The distributions were fit as full normal distributions with means above and below zero; to compare with a single-sided exponential model, the absolute value of the mean was used. The new model with positive mean was folded at $x = 0$ to capture the wrapped behavior of the absolute difference. Given mean μ_{ij} and variance σ_{ij}^2, the distribution of disagreeing rates can be approximated as:

$$p_d^{ij}(x) = \begin{cases} f_n(x; \mu_{ij}, \sigma_{ij}) + f_n(-x; \mu_{ij}, \sigma_{ij}), & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{5.4}$$

where f_n is the normal probability density function.

With shapes for both functions and parameters for the disagreeing portion, the exact structure final probability distribution can be described:

$$p^{ij}(x) = \alpha_{ij} p_a^{ij}(x) + (1 - \alpha_{ij}) p_d^{ij}(x) \tag{5.5}$$

$$= \alpha_{ij} \lambda_{ij} e^{-\lambda_{ij} x} + (1 - \alpha_{ij})(f_n(x; \mu_{ij}, \sigma_{ij}) + f_n(-x; \mu_{ij}, \sigma_{ij})) \tag{5.6}$$

for non-negative x. The remaining task is to derive α_{ij} and λ, which was again accomplished using MLE. For these analyses, the *mle* function in Matlab was used with flags for probability optimization. This resulted in a complete distribution that fit the data very well (see Figure 5.1c).

The final portion of the closeness is the derivation of cutoffs, enabled by the modeling of contributing distributions. Error probability for the cutoff $c_{i,j}$ is the cumulative distribution over the complementary domain:

$$P(error|agree) = P(error|disagree) \tag{5.7}$$

$$\alpha_{ij} \int_{c_{i,j}}^{\infty} \lambda_{ij} e^{-\lambda_{ij} u}\, du = (1-\alpha_{ij}) \int_{0}^{c_{i,j}} [f_n(x;\mu_{ij},\sigma_{ij}) + f_n(-x;\mu_{ij},\sigma_{ij})]\, du \tag{5.8}$$

$$\alpha_{ij}(1 - F_a^{ij}(c_{i,j})) = (1-\alpha_{ij}) F_d^{ij}(c_{i,j}) \tag{5.9}$$

where $F(x)$ is the cumulative distribution function for $p(x)$. This was solved using the *normcdf*, *expcdf*, and *fzero* functions for computing cumulative distribution functions and optimizing for the solution with numerical methods (see Figure 5.1d). The above process was repeated for each of the 630 unique pairs of different sensors.

5.2.2 Clustering Close Pairs into a True Rate

The cutoffs found in the previous section can be used to label each pair of sensors at a given timepoint to be classified as agreeing or disagreeing. It remains to be demonstrated how to find a single cluster of agreement from the pairs of rates. Consider an unweighted graph of the stream data, where each vertex v_i represents a stream rate estimate. Vertices v_i and v_j are connected with an edge e_{ij} if sensors i and j agree (see Figure 5.2).

Graph clustering literature is similarly aimed at multiple partitions where all clusters are equally valid and interesting. The classic graph interconnectivity problem for isolating a single grouping is the maximum clique, defined as the largest group of vertices where every node shares an edge with every other node. Finding a max clique, however, is an NP-Hard problem, known to be difficult for large graphs with many nodes due to the exponential growth of the solution set relative to the input size. Fast, exact algorithms for finding a max clique exist for small graphs [120], but finding the full list of max cliques can be time-consuming. Finding a single maximal clique with a high edge density requires checking many potential subgraphs before deciding that no larger clique exists, which can take multiple minutes (see Figure 5.3). Cases above 30 edges per node were rare, but they would have

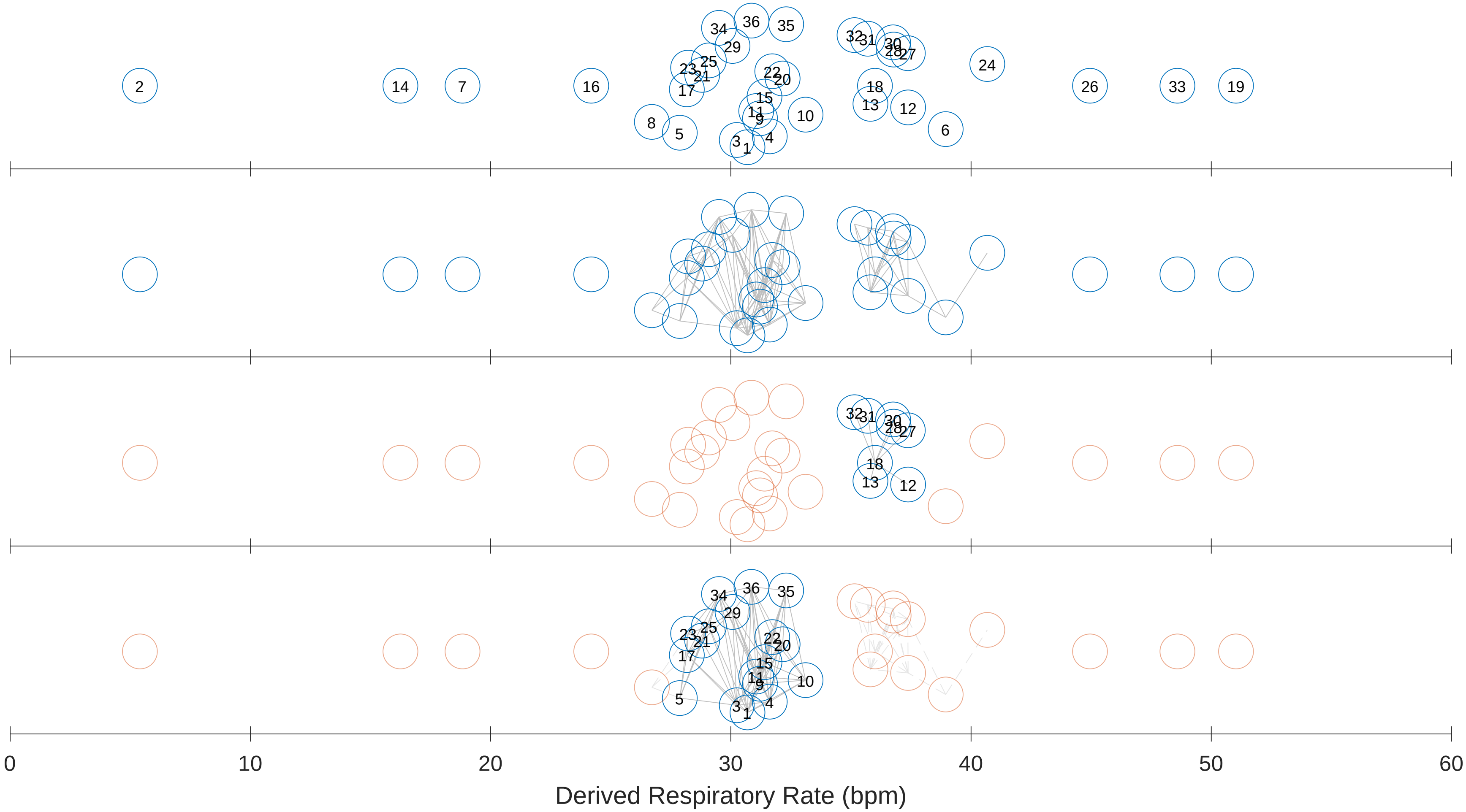


Figure 5.2: The process of clustering pairwise rate relationships into a single group of reliable means: (a) calculate 36 rates; (b) identify edges; (c) calculate edges per node, edges from node 18 shown here; (d) use all rates of largest according to tie-breaking rules. Here, all neighbors of rates 1 and 3 are included. (Vertical scaling for visibility only)

taken hours for every run of the clustering process on the data set. The maximal clique is not the ideal choice for a lightweight and easily repeatable metric of interconnectivity.

A more computable alternative is to group all vertices adjacent to the node with the most edges. This simpler statistic closely tracks the max clique size while minimizing the computation needed to find tying sets. Tying edge counts are often the product of a single cluster, with either the same or very similar groups of constituent nodes. On the other hand, some signal groups produce two substantially different clusters of equal size. To break ties, the previous set of used points is maintained; all nodes with the maximum edge count that appeared in the previous set of nodes are considered equally good. The tie-breaking set of used nodes includes all rates adjacent to all reappearing max nodes. In the event that no max nodes appeared in the previous solution set, all max nodes and their adjacencies are considered to be useful. The previous node set is initialized as the empty set, so all max nodes are included on the initial timestep.

With the agreeing cluster computed for each timepoint, the true-rate mean can be calculated as the average of agreeing rates. Figures 5.4 through 5.9 demonstrate the accuracy of the kinematics-derived rate against the flow standard. The technique reliably finds the respiratory rate curve in a variety of stable and unstable breathing conditions.

5.3 Evaluating True-Rate Mean Signals

The following analysis was performed with the benefit of having both a laboratory data set with spirometric flow rate data as well as a clinical ED data set with notable motion noise in many segments. The noisy clinical data were used for the determination of closeness thresholds, while the cleaner laboratory data provided a gold-standard reference. The laboratory setup included a bilateral SCM sensor pair, while the ED setup used only a single SCM sensor. To compensate, the left and right SCM resampled rate series for each stream were averaged together to provide a single set of six streams for all sensor pair-related algorithms.

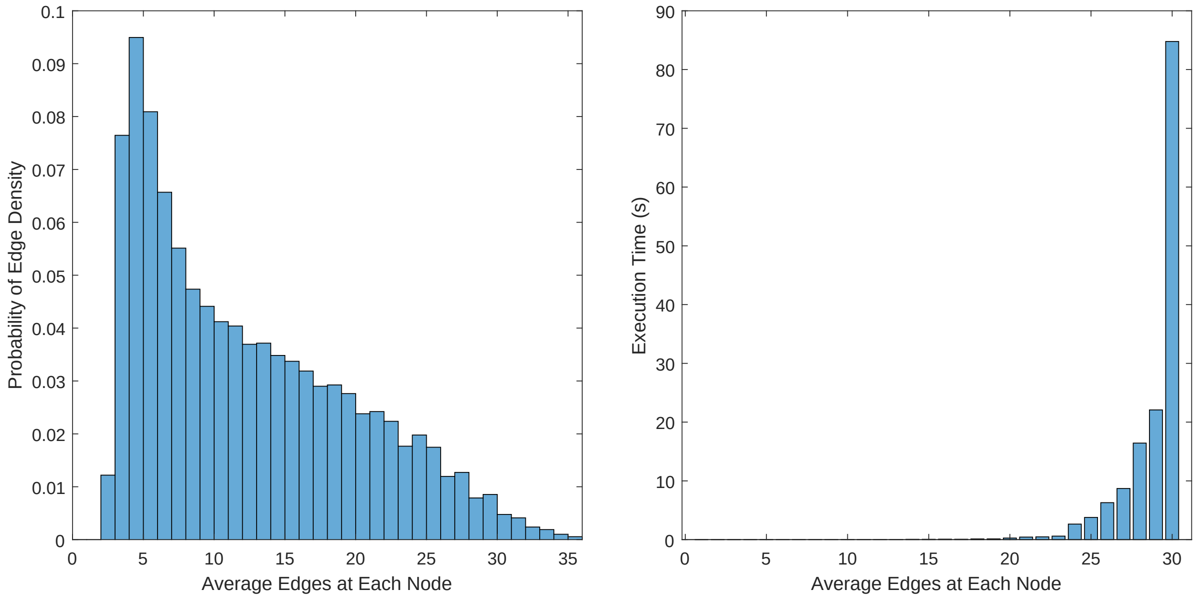


Figure 5.3: Edge Density Implications: A comparison of (a) average edge density of pairwise difference graphs, and (b) the time required to count the number of max cliques in a graph of varying edge density using a fast technique. Finding maximal cliques reaches a fundamental limit in efficiency when graphs are very dense. Though these events are relatively rarely in our data set, it is enough of an issue that would slow processing in real-time- and post-processing and could be unnecessarily costly in situations where the answer is clear due to high agreement.

5.3.1 Computing Cutoffs

More than 79,000 rate samples were derived for each of 36 sensor streams on 108 ARK patient recordings. The data set contained a mix of collections that were supervised by a research team member and unsupervised with an audio reminder and visual indication from sensor LEDs. Naturally, the supervised collections were much cleaner on the whole, but the inclusion of noisier segments allowed for more realistic distributions to be visualized, analyzed, and sampled for further modeling.

To derive cutoffs, the above process for modeling distributions was applied to each pair of sensor streams. The related distributions were sampled by taking the absolute value of rate differences between each pair of sensors. Rates for each sensor stream were then randomly permuted; unrelated distributions were derived by using 2000 rates from a pair of distributions, without replacement, and taking the difference between them. These were

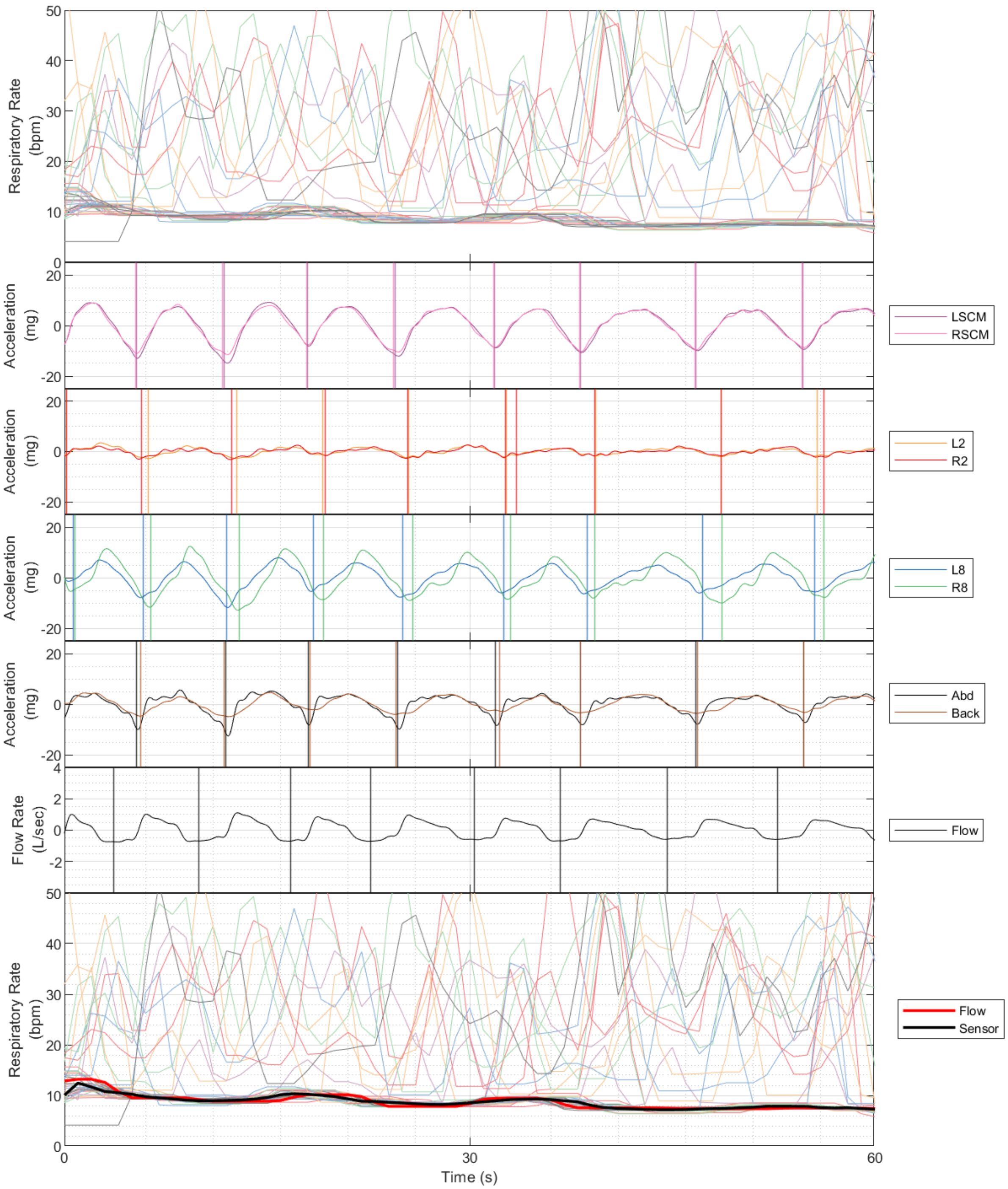


Figure 5.4: Steady, low-rate breathing, with instant rate estimates from the 36 streams (top), z-axis accelerometer and flow rate recordings (middle), and rate estimates with instant rate from the cluster mean and flow signals (bottom).

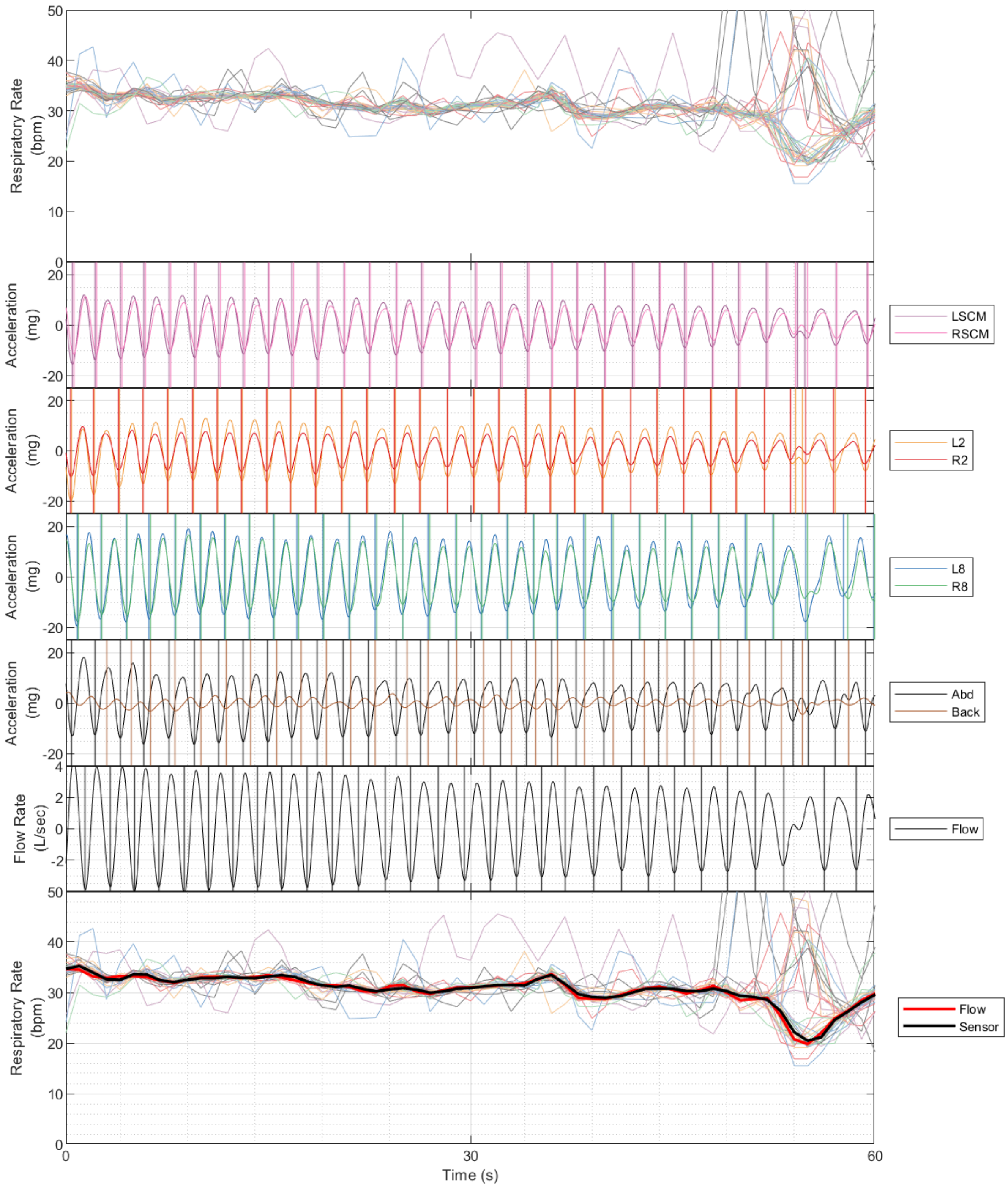


Figure 5.5: Steady, high-rate breathing, with instant rate estimates from the 36 streams (top), z-axis accelerometer and flow rate recordings (middle), and rate estimates with instant rate from the cluster mean and flow signals (bottom).

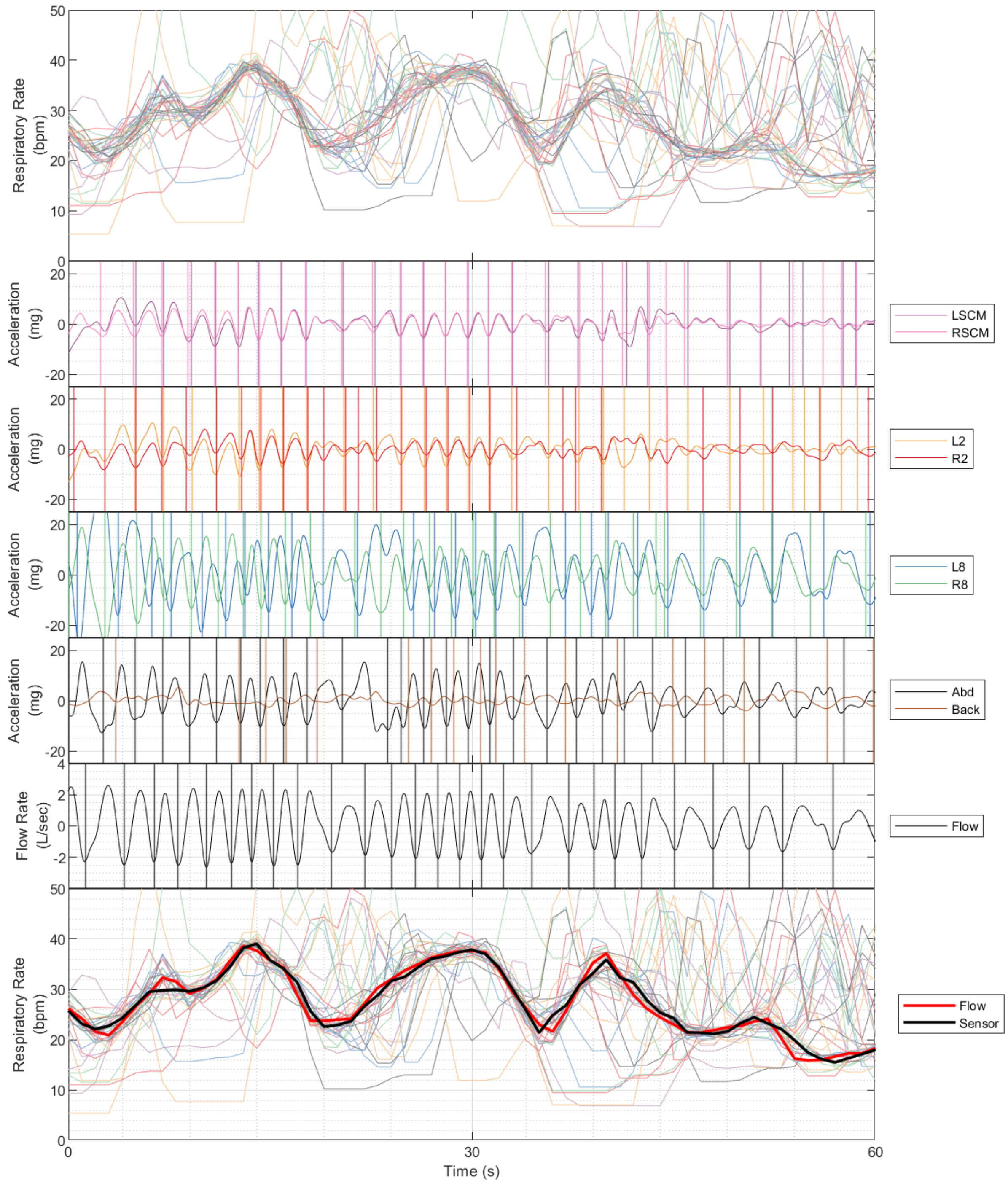


Figure 5.6: Erratic breathing, with instant rate estimates from the 36 streams (top), z-axis accelerometer and flow rate recordings (middle), and rate estimates with instant rate from the cluster mean and flow signals (bottom).

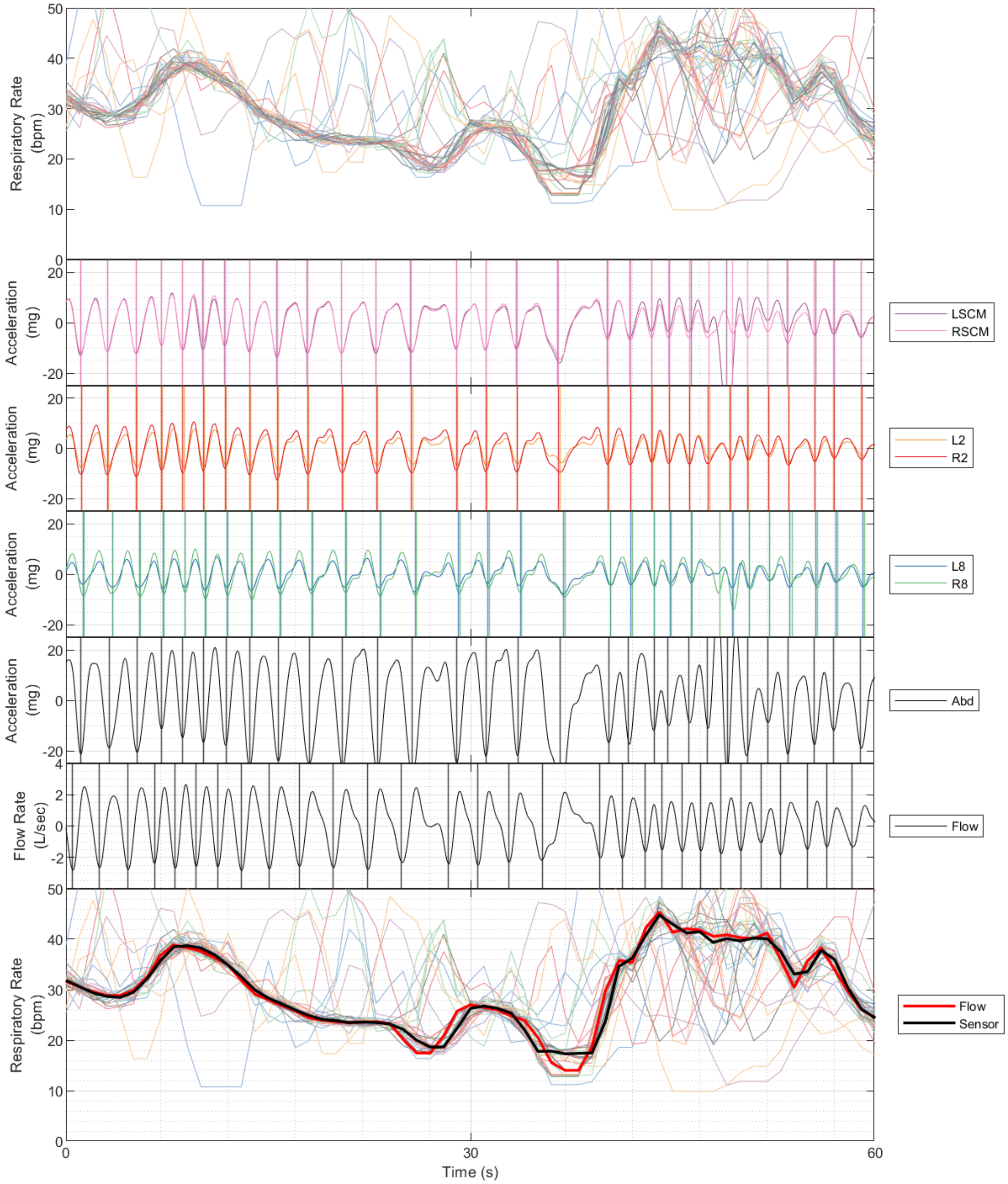


Figure 5.7: Erratic breathing with non-stationary reacceleration, with instant rate estimates from the 36 streams (top), z-axis accelerometer and flow rate recordings (middle), and rate estimates with instant rate from the cluster mean and flow signals (bottom).

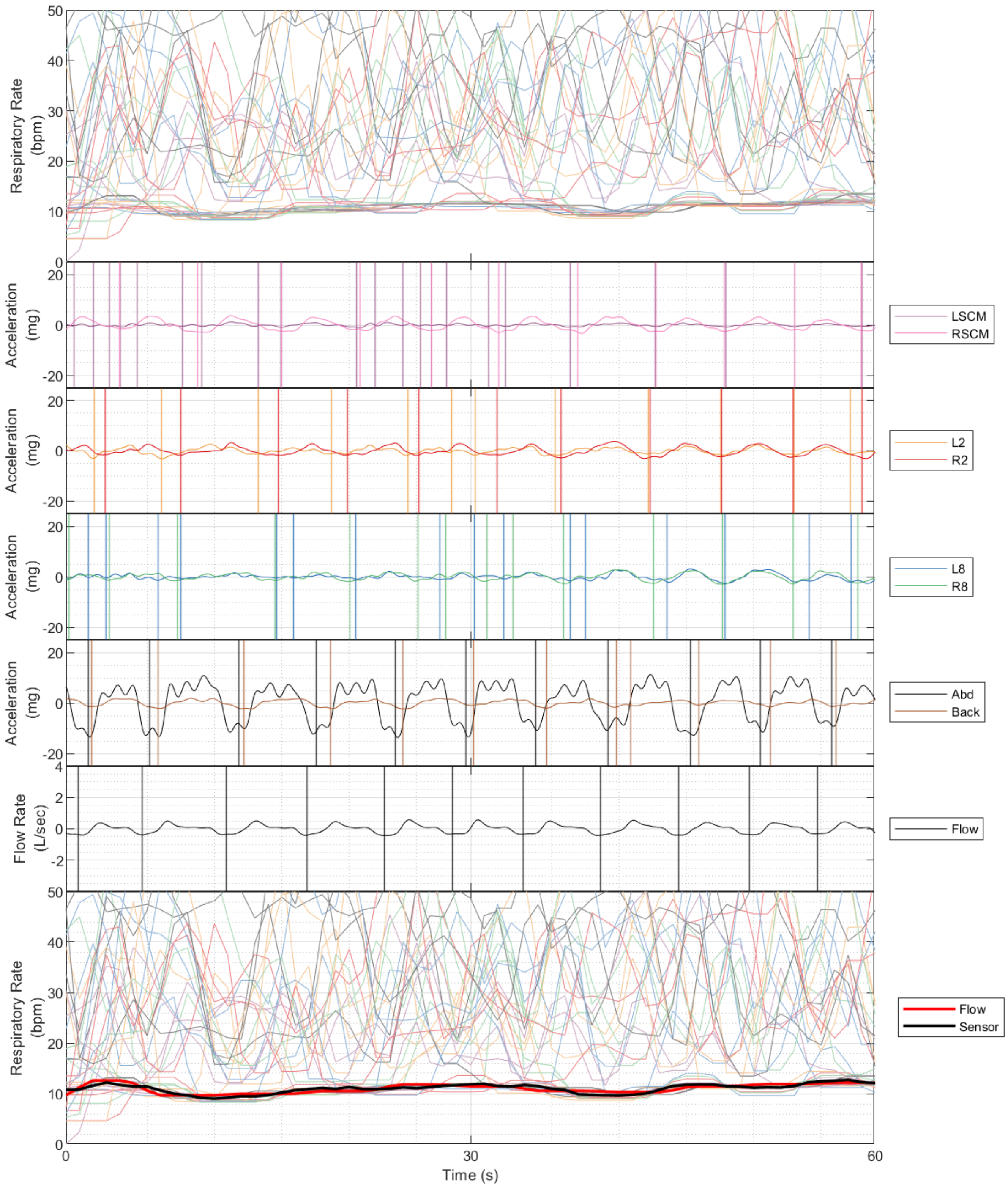


Figure 5.8: Stable breathing with noisy collection of rates, with instant rate estimates from the 36 streams (top), z-axis accelerometer and flow rate recordings (middle), and rate estimates with instant rate from the cluster mean and flow signals (bottom).

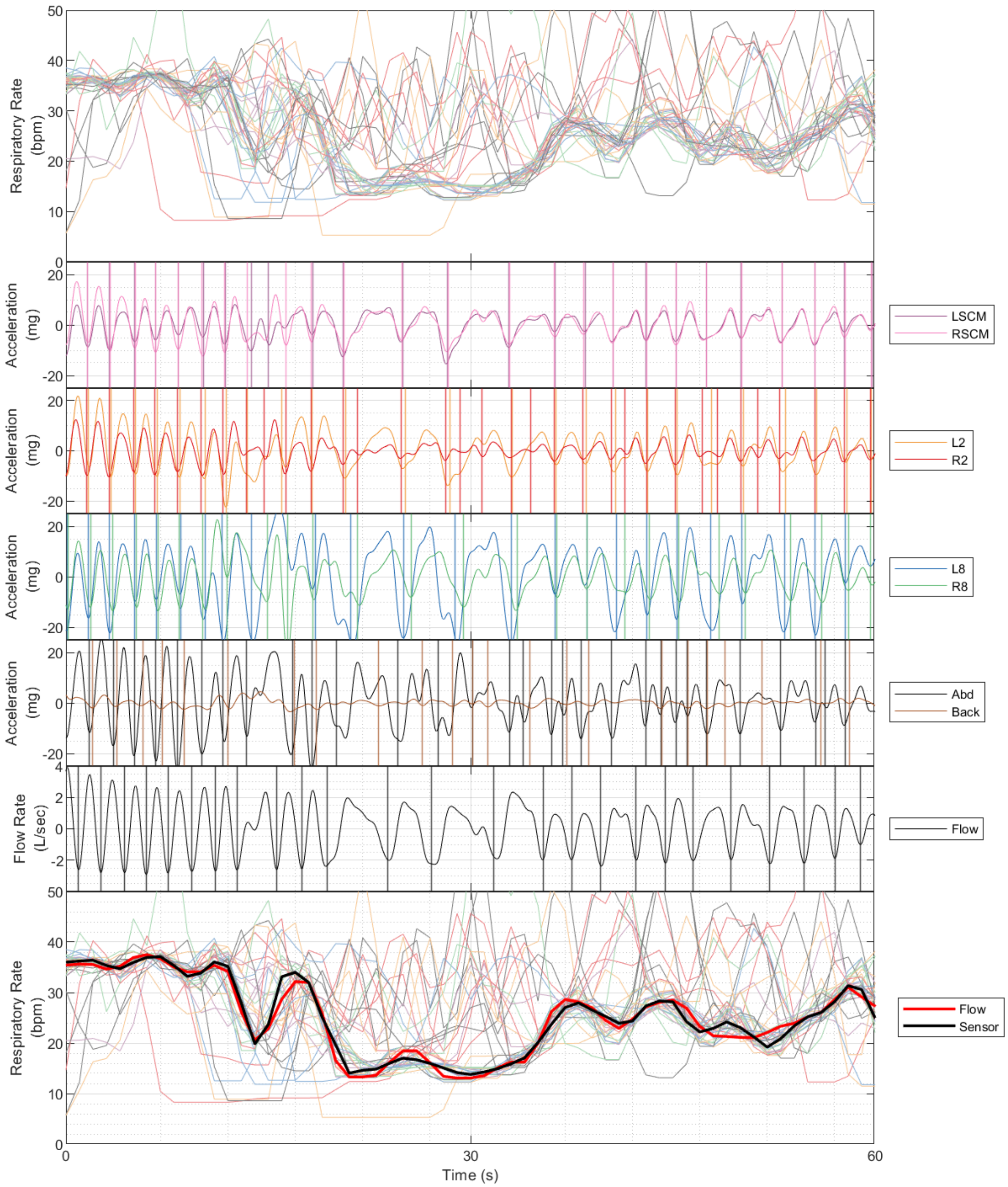


Figure 5.9: Varying rate of breathing with noisy collection of rates, with instant rate estimates from the 36 streams (top), z-axis accelerometer and flow rate recordings (middle), and rate estimates with instant rate from the cluster mean and flow signals (bottom).

visualized, and appeared close to normal. Each unrelated pairwise difference distribution was fit with a normal curve. The normal fits were folded by taking the absolute value of the mean, restricting the support of the function to above zero, and adding the negative probability to the probability corresponding positive input. For the final step of fitting, distributions of absolute were desirable to simplify modeling with a single exponential. Truncation before fitting was attempted, but it failed to effectively capture the mean of the difference of sensors, which was evident from the non-equal means of individual stream rate distributions and the untruncated normal means differing from zero. Fitting was done on a per-pair basis due to the variation in means between sensor streams.

With the distributions in hand, the final likelihood estimation was performed to find the exponential constant and the interpolation factor between the exponential and truncated normal distributions. The difference of weighted CDFs was solved numerically, producing a cutoff that balanced the weighted probabilities of error.

Tables 5.1 through 5.10 below document the fitted distribution parameters and the resulting cutoffs derived from the noisy data, aggregated by sensor and by stream. The tables justify the individual fitting process; while sensors performed very similarly, differences in the behavior of rates of stream types became evident through the fitting process. Notably, the mean rates from gyroscopes were substantially higher, causing much larger means for unrelated pairwise differences. When factored into the full distribution, exponential means were lower (less steep) and the interpolation favored the noisy distribution more heavily. Resulting cutoffs are lower for stream pairs where one or both of the streams is a gyroscope.

5.3.2 Measuring Rate with ARK

The modeling process produced cutoffs for each pair of the 36 sensor streams. The clustering process described above was applied to the EPCL data set. The results are striking when compared to individual sensor or naive mean results from Chapter 4. The coefficient of

Sensor	SCM	L_2	R_2	L_8	R_8	Abd
SCM	5.71	4.64	4.68	4.49	4.55	4.95
L_2	4.64	5.19	4.34	4.31	4.44	4.74
R_2	4.68	4.34	5.21	4.29	4.35	4.65
L_8	4.49	4.31	4.29	4.84	4.14	4.58
R_8	4.55	4.44	4.35	4.14	4.92	4.45
Abd	4.95	4.74	4.65	4.58	4.45	5.39

Table 5.1: Normal Fit Means (bpm), Average by Sensor

Sensor	SCM	L_2	R_2	L_8	R_8	Abd
SCM	15.46	15.22	15.20	15.16	15.14	15.52
L_2	15.22	14.91	15.04	14.97	14.92	15.34
R_2	15.20	15.04	14.87	14.90	14.83	15.27
L_8	15.16	14.97	14.90	14.89	14.79	15.27
R_8	15.14	14.92	14.83	14.79	14.67	15.20
Abd	15.52	15.34	15.27	15.27	15.20	15.63

Table 5.2: Normal Fit Standard Deviations (bpm), Average by Sensor

Sensor	SCM	L_2	R_2	L_8	R_8	Abd
SCM	1.10	1.07	1.08	1.12	1.09	1.18
L_2	1.07	0.99	1.06	1.10	1.07	1.12
R_2	1.08	1.06	1.02	1.09	1.06	1.11
L_8	1.12	1.10	1.09	1.02	1.05	1.12
R_8	1.09	1.07	1.06	1.05	1.02	1.08
Abd	1.18	1.12	1.11	1.12	1.08	1.09

Table 5.3: Exponential Distribution Means (bpm), Average by Sensor

Sensor	SCM	L_2	R_2	L_8	R_8	Abd
SCM	0.31	0.32	0.33	0.32	0.32	0.29
L_2	0.32	0.35	0.34	0.34	0.34	0.30
R_2	0.33	0.34	0.36	0.35	0.35	0.31
L_8	0.32	0.34	0.35	0.41	0.39	0.35
R_8	0.32	0.34	0.35	0.39	0.39	0.34
Abd	0.29	0.30	0.31	0.35	0.34	0.34

Table 5.4: Interpolation Factors, Average by Sensor

Sensor	SCM	L_2	R_2	L_8	R_8	Abd
SCM	1.82	1.78	1.82	1.85	1.81	1.84
L_2	1.78	1.79	1.84	1.90	1.85	1.81
R_2	1.82	1.84	1.85	1.90	1.87	1.82
L_8	1.85	1.90	1.90	1.97	1.96	1.96
R_8	1.81	1.85	1.87	1.96	1.94	1.87
Abd	1.84	1.81	1.82	1.96	1.87	1.88

Table 5.5: Pairwise Difference Cutoffs, Average by Sensor

Stream	a_x	a_y	a_z	g_x	g_y	g_z
a_x	0.54	0.98	1.88	8.50	8.21	8.17
a_y	0.98	0.94	2.54	9.16	9.00	8.82
a_z	1.88	2.54	1.41	6.62	6.47	6.33
g_x	8.50	9.16	6.62	0.92	0.68	0.72
g_y	8.21	9.00	6.47	0.68	0.57	0.48
g_z	8.17	8.82	6.33	0.72	0.48	0.52

Table 5.6: Normal Fit Means (bpm), Average by Stream

Stream	a_x	a_y	a_z	g_x	g_y	g_z
a_x	14.06	13.98	14.85	14.87	14.66	14.94
a_y	13.98	14.02	14.81	14.84	14.69	15.01
a_z	14.85	14.81	15.59	15.70	15.48	15.73
g_x	14.87	14.84	15.70	15.81	15.55	15.88
g_y	14.66	14.69	15.48	15.55	15.56	15.61
g_z	14.94	15.01	15.73	15.88	15.61	15.79

Table 5.7: Normal Fit Standard Deviations (bpm), Average by Stream

Stream	a_x	a_y	a_z	g_x	g_y	g_z
a_x	1.21	1.41	1.13	1.08	1.04	1.10
a_y	1.41	1.38	1.31	1.06	1.05	1.11
a_z	1.13	1.31	1.10	1.01	1.00	1.04
g_x	1.08	1.06	1.01	1.00	0.97	1.01
g_y	1.04	1.05	1.00	0.97	0.92	0.95
g_z	1.10	1.11	1.04	1.01	0.95	0.98

Table 5.8: Exponential Distribution Means (bpm), Average by Stream

Stream	a_x	a_y	a_z	g_x	g_y	g_z
a_x	0.47	0.45	0.45	0.30	0.33	0.30
a_y	0.45	0.46	0.42	0.29	0.28	0.28
a_z	0.45	0.42	0.41	0.29	0.31	0.29
g_x	0.30	0.29	0.29	0.32	0.32	0.31
g_y	0.33	0.28	0.31	0.32	0.34	0.31
g_z	0.30	0.28	0.29	0.31	0.31	0.31

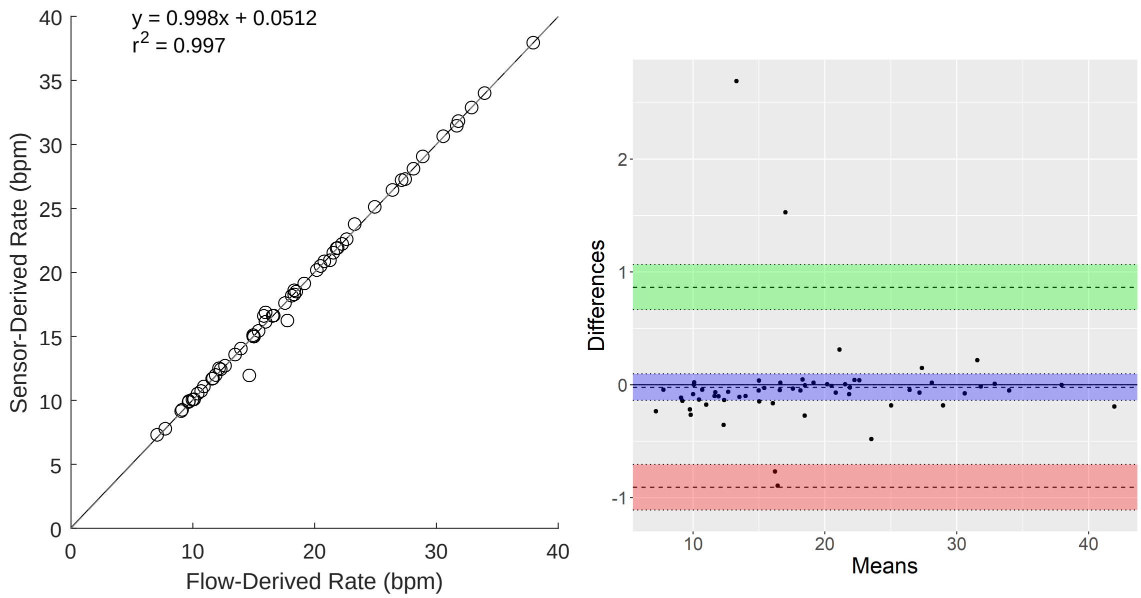
Table 5.9: Interpolation Factors, Average by Stream

Stream	a_x	a_y	a_z	g_x	g_y	g_z
a_x	2.31	2.47	2.16	1.79	1.82	1.81
a_y	2.47	2.47	2.31	1.74	1.72	1.76
a_z	2.16	2.31	2.09	1.69	1.72	1.70
g_x	1.79	1.74	1.69	1.70	1.65	1.67
g_y	1.82	1.72	1.72	1.65	1.64	1.59
g_z	1.81	1.76	1.70	1.67	1.59	1.65

Table 5.10: Pairwise Difference Cutoffs, Average by Stream

determination of the regression line between flow- and sensor-derived rate means improved to 0.997 (see Figure 5.10a), indicating a very close fit between the two mean values. The Bland-Altman analysis confirmed the stability of residuals, with a bias of 0.05 bpm and 95% limits of agreement at 0.88 bpm (see Figure 5.10b).

Pairing with the encouraging mean-rate results, the time-series errors improved over both the naive mean and best-sensor approaches outlined in the previous chapter. Errors dropped to almost half from the original best results; the method achieves 3.7% error on average, with a minimum error of 0.76% (see Figure 5.12). The average RMS error was 0.89 bpm, with a best of 0.21 bpm. These results indicate that the method produces a mean time-series that is consistent with reference signals in spite of local variations in individual sensors, previously shown to be inaccurate. Combined with the regression and Bland-Altman analysis, this evaluation demonstrates that the ARK system is an effective tool for studying rate dynamics. Additionally, the information contained in the clustered streams provides evidence of the prevalence of motion noise artifacts, allowing analysis and removal of noise from body movement and coughing. The remaining error can be attributed to the minor



(a) (b)

Figure 5.10: Correlation Analysis of Combinational Method: Means correlated better than with previous simpler combinations; Bland-Altman analysis gave a bias of 0.02 bpm (95% limits of agreement: 0.88 bpm).

lack of synchronization between the air flow and the kinematics of the chest, as well as uncertainty due to the re-synchronization of the flow data to the sensor data.

The instantaneous error was low enough that measurements of variance should hold meaning. To test this hypothesis, a similar regression and Bland-Altman analysis was performed (see Figure 5.11). The regression had a coefficient of determination of 0.95, while the Bland-Altman showed a low bias (0.02 bpm) and high agreement (95% limits of agreement: 0.79 bpm). This indicates that the variance of the series is also well preserved with the Hilbert-Berger cluster technique on the kinematic sensor data.

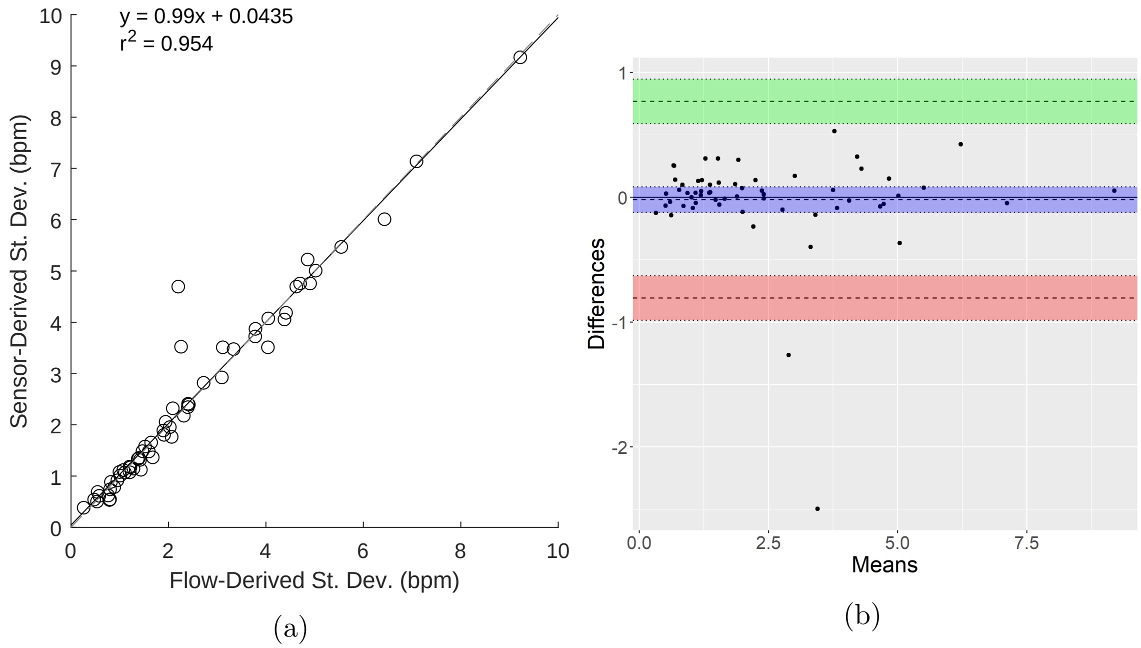


Figure 5.11: Correlation Analysis of Combinational Method: Variance of the time series also correlated well, with some noisier outliers; Bland-Altman analysis gave a bias of 0.02 bpm (95% limits of agreement: 0.79 bpm).

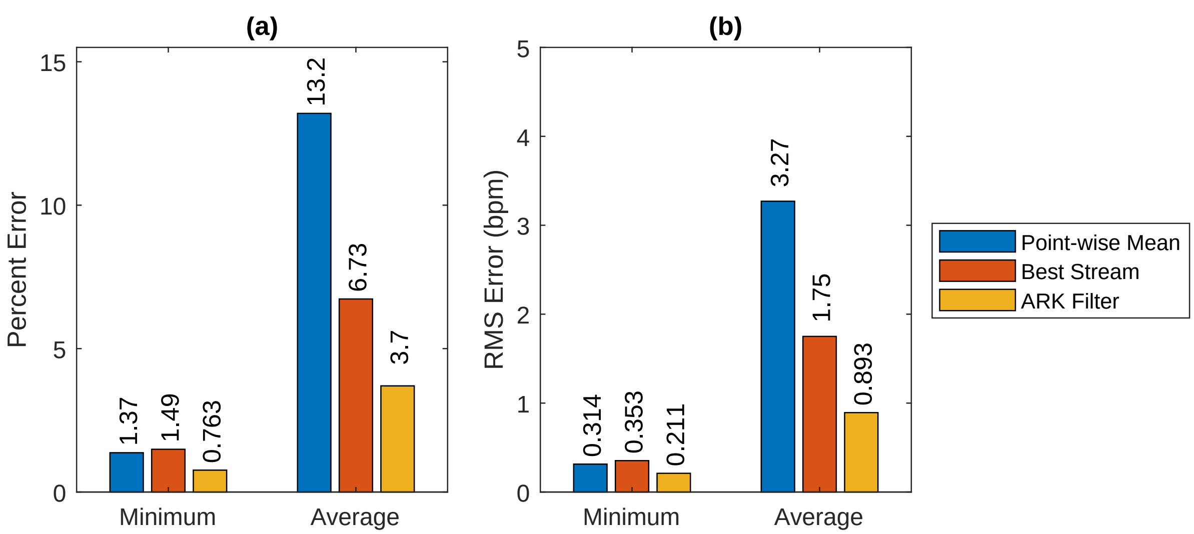


Figure 5.12: Comparison of time-series errors for combinational method. (a) shows percent error metrics, while (b) shows the RMS error versions. The robust implementation achieved just over half the average error of the best individual sensor, and similarly improved on the minimum error mark of the naive average of three sensor streams.

Chapter 6

Multi-Modal Analysis of ARK Signals for Indications of Respiratory Distress

The previous chapter presented a method for identifying a true-rate mean signal out of noise, but a robust implementation must account for the possibility of no useful signal. Additionally, analyses of other aspects of breathing like the variation of true rate and contributions of compartments require knowledge of segments of usable data. Without some confidence that the data used contain signals of interest, secondary metrics would be overwhelmed with noise that would distort means and exacerbate variance. A filter is presented below using knowledge of the agreeing rates to find areas of high agreement; metrics are extracted using the filtered data that begin to quantify respiratory distress. Preliminary statistical analysis demonstrates predictive power on a set of metrics exclusively derived from an ARK system.

6.1 Filtering the Mean Signal

In order to find meaningful data in the new mean signal, the noisy segments must be removed and the clean segments kept. One approach would be to label a data set for each timepoint as either noise clean and build a more traditional classifier; this is easier said than done as there are 36 signals and one hundred intervals per two minute collection. Labelling would again be extremely time intensive and prone to human misjudgement without specific comments from a physician taken during the sample. There are too many data to consider while making a reasoned, manual analysis. The semi-psychosomatic nature of breathing complicates the problem even further: a patient could be less likely to exhibit a gross motion effect under observation, complicating attempts to collect a physician-labelled data set of errors for filtering purposes. Given our collection style and the pressing need for a system to find usable data, the problem calls for another automated solution.

An intuitive approach to finding good data is to the use the number of rates used at a given time as a measure of similarity of the recorded rates. Two methods were explored: (1) a more liberal inclusion criteria to improve recruitment of data samples, and (2) a more conservative approach with higher guarantee of the quality of data. The liberal algorithm operated on the principle of saving roughly 50% of data, keeping more when the data are high quality and less when particularly low quality. At 100 samples, 50 samples would render moments and other common statistics valid from the perspective of sample size,and it would provide a decently sampled signal for signal processing methods. High and low quality were determined by assessing the distribution of rate counts as a comparison point to the series of rate counts for each two-minute segment. The median was selected as the metric of interest for two-minute rate counts; given that exactly half the data are above and below, the median serves as a proper bellwether for the distribution. Averages would be skewed heavily by outliers, resulting in unnecessarily high or low cutoffs and ruining the selectivity of the filter. A high median was defined as 18 rates and a low median defined as 12 rates. 18 is well situated at half of the total, indicating a clear collection of rates. Additionally, earlier iterations of cutoffs produced median and lower quartile rates used at 18 and 12, inspiring the two-threshold filter. For high median segments, all mean rate samples with 18 or more rates used were included, while all samples with 12 or less rates were excluded for low median segments. For data with medians between these two points, all data above the median were considered valid. This set of rules applied a sliding scale: keeping reasonably good data, rejecting bad data, and defaulting to the median cutoff when data fall in between.

While this method was built as a reasonable approach, the filter suffered due to the more relaxed guidelines. While all low rate counts were thrown out, the data just above the threshold were too noisy. For instance, all included data from a segment with medians counts at 13 and max counts at 15 would contain only points where more rates disagreed than not, yet half the data would still be counted. While this was somewhat effective at finding the true rate, it introduced substantial amounts of noise.

An alternative approach would be to establish a single threshold for rate counts. As opposed to the relative approach, the specificity of the absolute classification is a free parameter; noise events can be eliminated to whatever degree desired. Still, increasing the quality of the selected data comes at the price of reducing the total available data for analysis. A balance between sensitivity and specificity is needed.

To find the absolute cutoff for rate counts, filter results were visualized for thresholds ranging from 50% to 90%. A cutoff at 50% (18 rates) was chosen for three reasons: 50% produced higher sensitivity to higher cutoffs, with little introduced noise, and with a resulting signal that comprised half of the total streams. The chance that a cluster using at least half the rates is incorrectly derived is small for sufficiently high rates (>15 bpm), as practically any other cluster would have less agreement. For lower measured rates, high agreement can be caused by sufficiently apneic or noise-biased signal that no longer generates clean periodic waves. The lack of periodicity can be measured as extended intervals of repetition, resulting in a low rate estimate. Given that both apnea and gross motion are likely to affect the whole torso, false low rates can be detected with high agreement. This failure mode of similarity agreement is strong enough that even more stringent absolute filters had issues rejecting the inaccurate estimates. Thus, an absolute cutoff at 50% keeps the maximum amount of data while still having good confidence that the measured signal is meaningful. Undoubtedly, more complex filters to handle noise and apnea before evaluating similarity would be useful, but the 50% mark is sufficiently robust for this analysis.

The result is an unevenly sampled series of instantaneous respiratory rates, but also a collection of intervals where component signals were judged to be synchronous and useful. While not every sensor stream is guaranteed to be meaningful at every point in time, it's a much stronger assumption than using all data points and hoping patterns emerge from noise.

Figures 6.1 through 6.6 demonstrate the filtering capability of separating good data from transient and steady noise in the recordings. The technique still reliably finds the respiratory rate curve in a variety of stable and unstable breathing conditions despite the noise.

6.2 Deriving Metrics

6.2.1 Mean Rate

Given the filtered mean signal, the most obvious metric to compute is the mean-of-means. This serves as an accurate representation of the two-minute average respiratory rate. Average respiratory rate is not a perfect metric, as healthy and sick ranges overlap even for one individual, let alone a population. Still, this mean-of-means rate is the careful aggregation of non-noise rate segments. This is preferable to the spectrum method, which bets on a clear, emergent peak, and peak-finding methods, which are not robust to the high-noise, widely-varying signals seen in respiratory kinematics. The rate filtering on the time-series contributing to the mean ensures that the final estimate is an accurate representation of two-minute segments.

6.2.2 Variance of Rate

The primary advantage of an accurate instantaneous respiratory rate time-series over averages is the evaluation of dynamic behavior. Again, respiratory rate variance is highly understudied; previous methods were not designed with the intent to measure variability

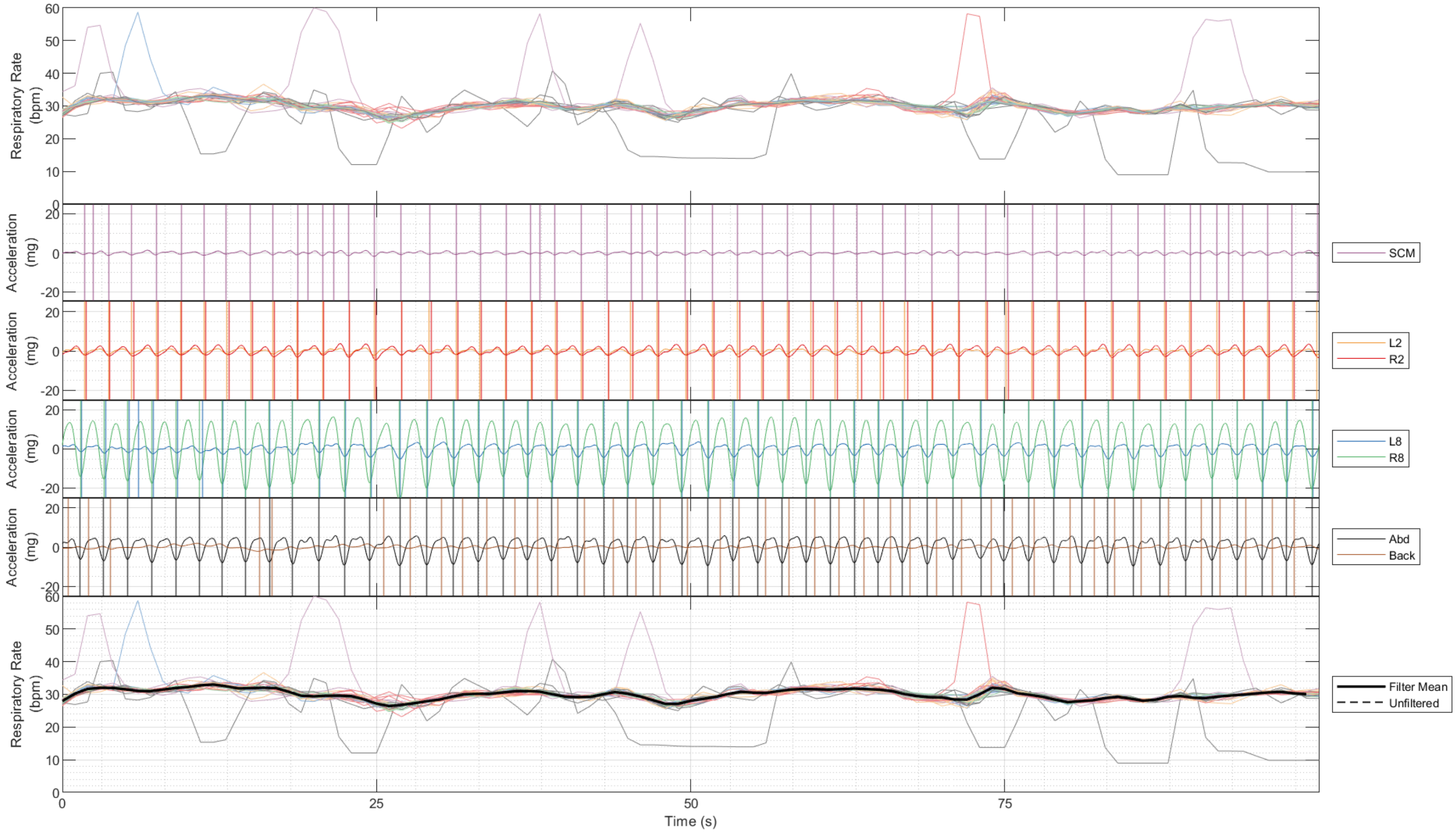


Figure 6.1: Clean, stable signal, with instant rate estimates from the 36 streams (top), z-axis accelerometer recordings (middle), and rate estimates with instant rate from the cluster mean signal (bottom).

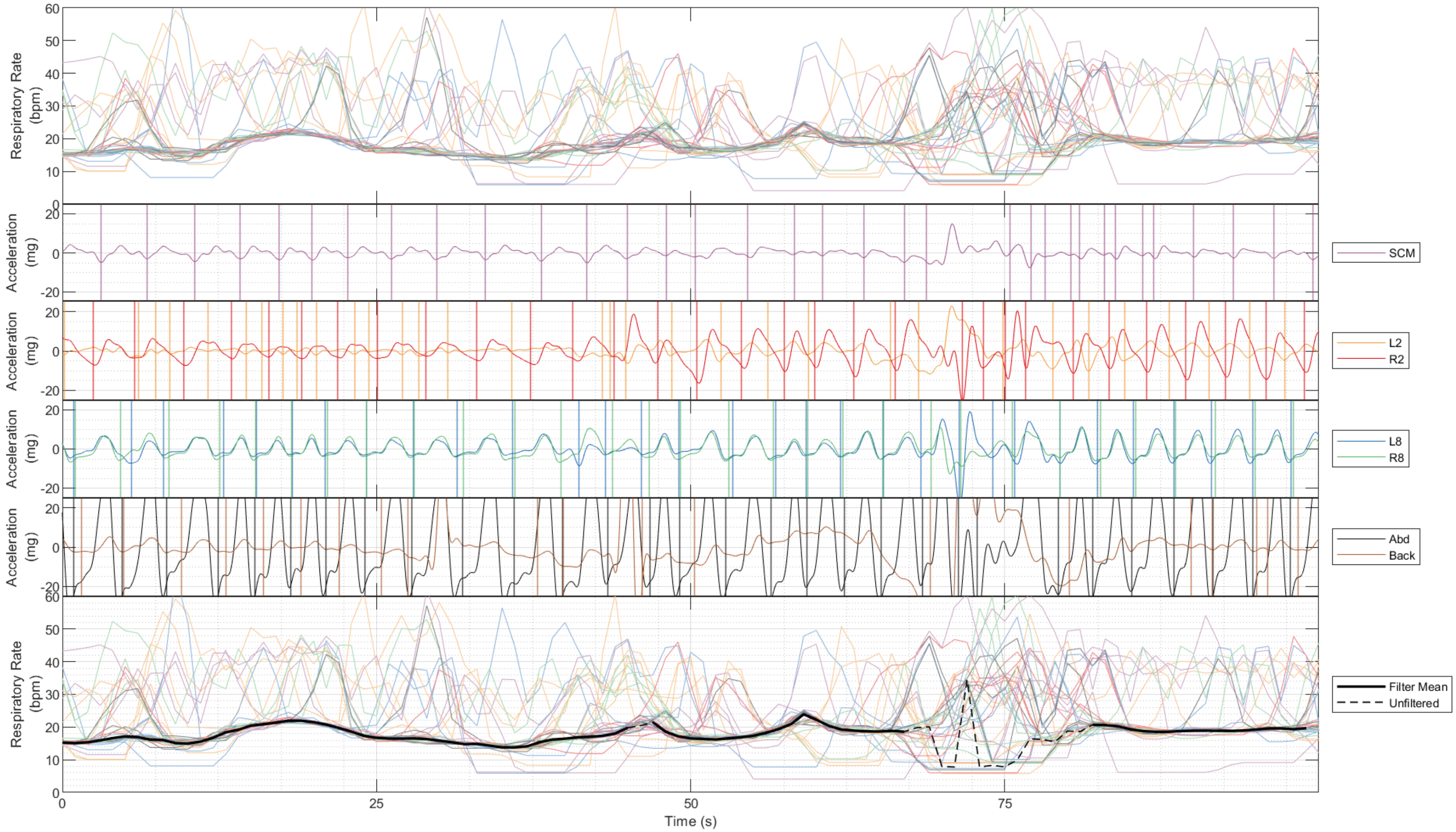


Figure 6.2: Stable signal with transient noise, with instant rate estimates from the 36 streams (top), z-axis accelerometer recordings (middle), and rate estimates with instant rate from the cluster mean signal (bottom).

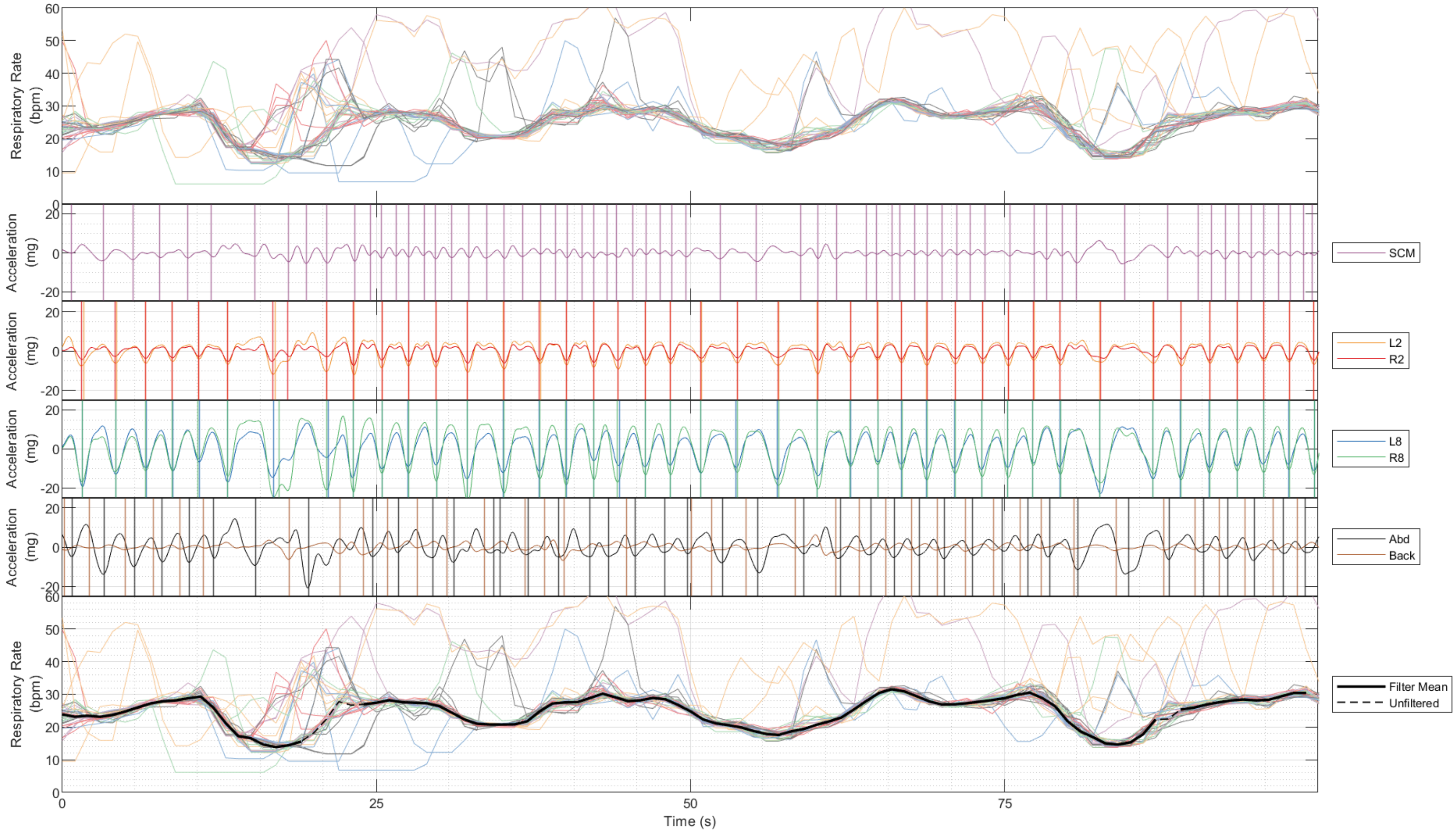


Figure 6.3: Periodic breathing pattern, with instant rate estimates from the 36 streams (top), z-axis accelerometer recordings (middle), and rate estimates with instant rate from the cluster mean signal (bottom).

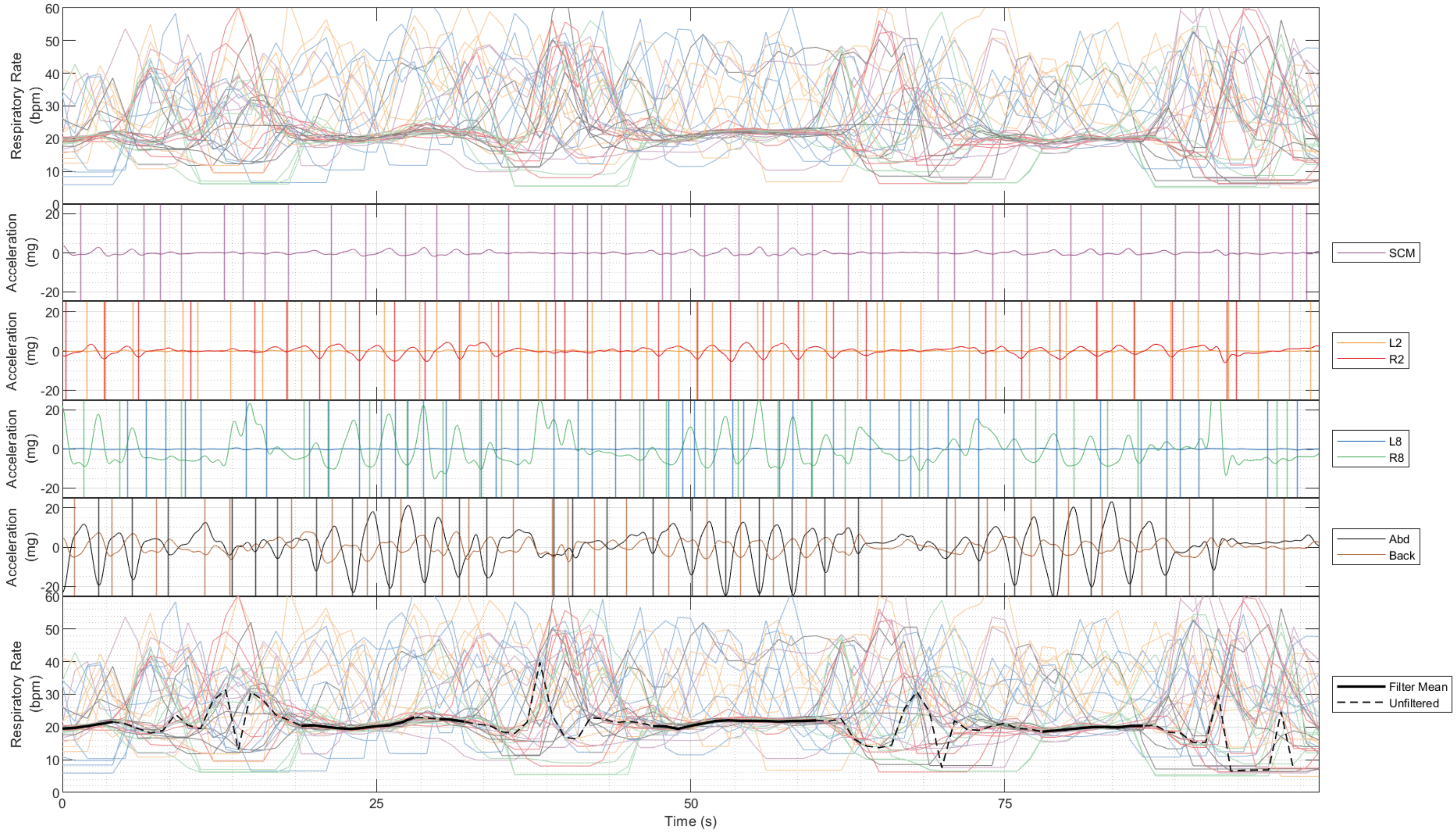


Figure 6.4: Cheyne-Stokes intermittent breathing pattern, with instant rate estimates from the 36 streams (top), z-axis accelerometer recordings (middle), and rate estimates with instant rate from the cluster mean signal (bottom).

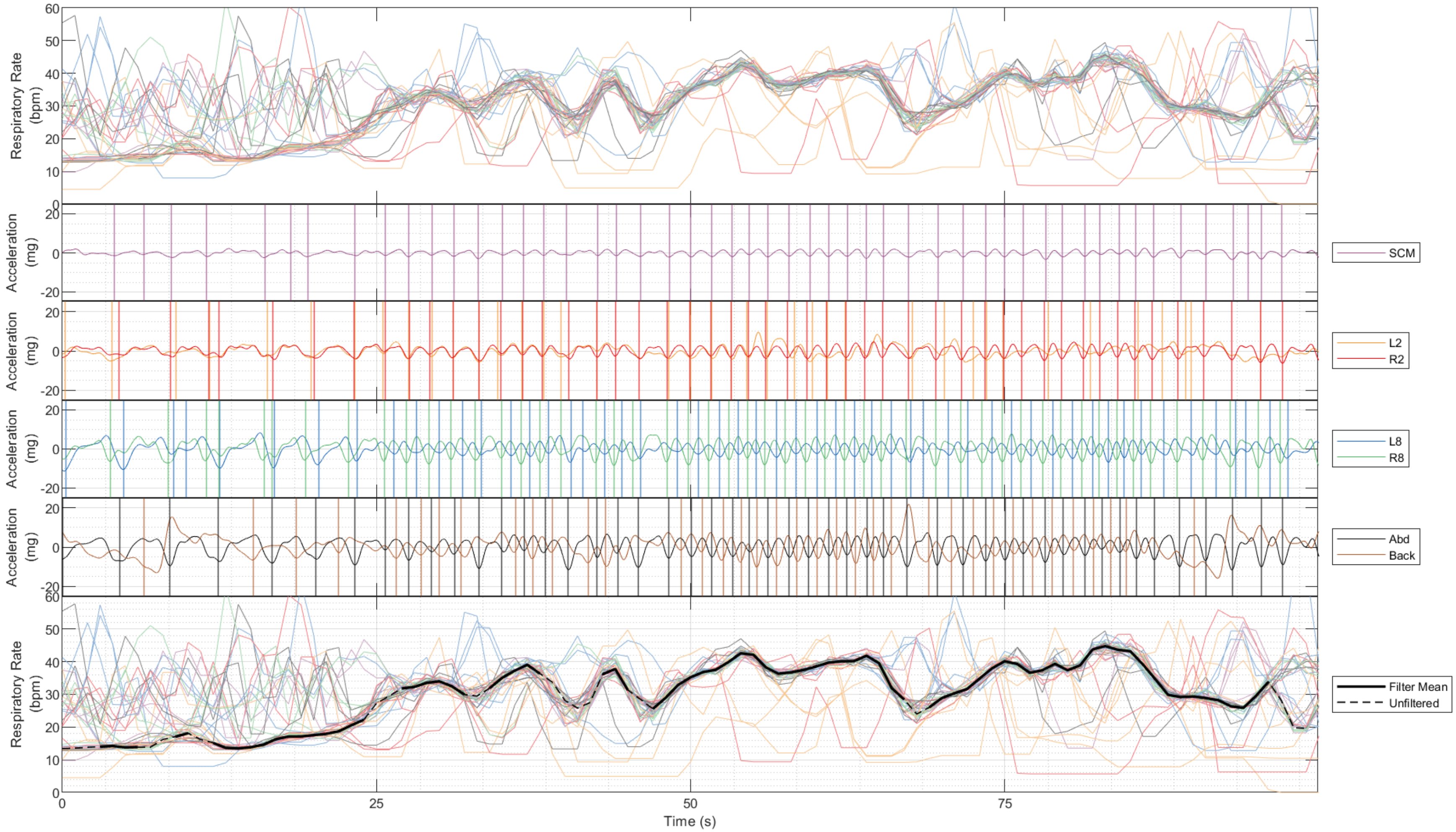


Figure 6.5: Unstable breathing rate, with instant rate estimates from the 36 streams (top), z-axis accelerometer recordings (middle), and rate estimates with instant rate from the cluster mean signal (bottom).

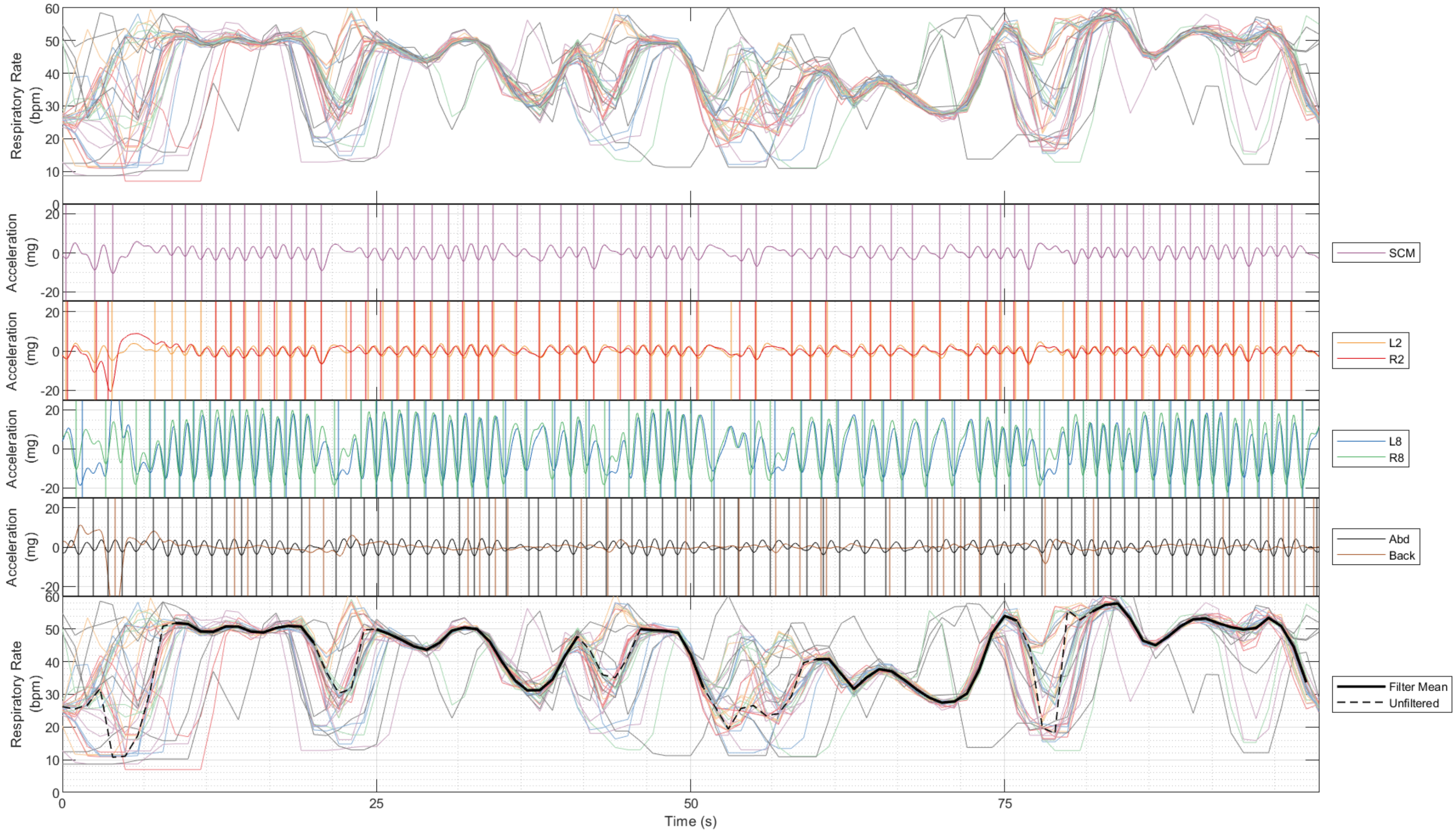


Figure 6.6: Highly unstable breathing rate, with instant rate estimates from the 36 streams (top), z-axis accelerometer recordings (middle), and rate estimates with instant rate from the cluster mean signal (bottom).

patterns. While it is difficult to prescribe how it could be used in medical contexts, similar measures for cardiovascular health have been proven to be effective in quantifying distress. For metrics, standard deviation is the classic metric for time-series variance, as with heart-rate variability. Due to the low-rate failure mode mentioned above, however, consistently high rates with low-rate noise can appear as highly-variant. To that end, a second metric set was examined, looking at first differences at different sample spacings (1, 2, and 3 seconds). The average, median, and maximum were calculated from the lists of differences of a two-minute collection. Both sets of metrics were included for further analysis.

6.2.3 Recruitment of Accessory Muscles

The textbook visual sign of respiratory distress is the recruitment of the scalene and SCM muscle groups around the base of the neck. Along with obvious changes like sudden increases in breathing rate (tachypnea), accessory muscle use is a surefire sign that the patient can no longer provide enough force with the primary muscle groups and is entering a state of distress. Currently, there does not exist a metric to quantify this aspect of labored breathing, despite the known clinical significance. While studies have been done with upper rib sensing (e.g. RIP, various inertial designs) and even with SCM sensing [72], but not in a way that is designed to make sense of the magnitude of motion during breathing. Breathing is a complex set of motions, the magnitudes are completely dependent on individual anatomy. Thus, previous studies have left magnitude analysis untouched due to the difficulty in drawing conclusions without analyses controlling for contributing factors.

For this analysis, magnitude of motion for a sensor was taken to be the Euclidean norm of the individual dimensions of the 3-axis acceleration signal. Each acceleration stream was processed for an envelope. While the amplitude function from the analytic representation provides an envelope, the wide frequency band of breathing signals introduces artifacts that are incorrectly expressed as oscillations in the amplitude at the same rate as breathing. The resulting envelope is too noisy for magnitude analysis without further smoothing. Still, the extrema of a signal are well-detected by the analytic representation. To avoid a non-robust smoothing procedure, the maxima and minima of each signal were combined into upper and lower envelopes using an uneven upsampling process. The original time vector was used as the target time vector; each time not containing a landmark was filled with the amplitude of the nearest landmark. This step wave was filtered such that the cutoff frequency was equal to the reciprocal of the longest time between landmarks. This guaranteed that frequencies appearing in the envelope accurately reconstructed while staying robust to potentially long gaps for noisy records. The method improves upon the stability of the analytic envelope while avoids the volatility of spline interpolation, which produces unusable results for signals with sparse landmarks due to low amplitude and even inverted envelopes where one swings past the other. The final acceleration envelopes ${}^{e}a_{sd}$ for each sensor s and dimension d were combined with a Euclidean norm into the sensor acceleration magnitude, ${}^{e}a_{s}$:

$$
{}^{e}a_{s}(t) = \sqrt{{}^{e}a_{sx}^{2}(t) + {}^{e}a_{sy}^{2}(t) + {}^{e}a_{sz}^{2}(t)} \tag{6.1}
$$

Comparisons between different sensors offer an alluring alternative to the less tractable absolute magnitude analysis. Enabled by the collection of data from multiple points on the chest, relative analysis can offer an indication of excessive accessory motion using the primary compartments as a reference. To this end, motion signals for the four vertical positions (SCM, upper rib, lower rib, and abdomen) were derived by combining bilateral signals with a simple averaging. While lateral differences in breathing motion are documented (i.e. flail chest), they are relatively unusual and almost always obvious to the naked eye. Thus, bilateral measurement was used to reinforce the reliability of measurements:

$$
{}^{e}a_{UR}(t) = \frac{{}^{e}a_{L2} + {}^{e}a_{R2}}{2} \tag{6.2}
$$

$$
{}^{e}a_{LR}(t) = \frac{{}^{e}a_{L8} + {}^{e}a_{R8}}{2} \tag{6.3}
$$

From the four regional signals, we calculated a ratio between the upper signals and lower signals, denoting this to be the index of Recruitment of Accessory Muscles (RAM). For the initial version, each pair of regions were combined with a simple average:

$$RAM(t) = \frac{{}^{e}a_{SCM}(t) + {}^{e}a_{UR}(t)}{{}^{e}a_{LR}(t) + {}^{e}a_{Abd}(t)} \tag{6.4}$$

While the metric displayed some discriminatory power, we hypothesized that the relationship could be improved by selecting signals to intentionally exacerbate the metric. By definition, RAM increases for larger upper compartment magnitudes and smaller lower compartment magnitudes. Given that different subjects may vary their recruitment of each component signal, the resulting upper and lower magnitudes may unnecessarily dilute the effect of upper muscle recruitment. To combat this effect, opportunistic RAM (opRAM) was developed as a companion metric. In contrast to RAM, opRAM is calculated for each 100Hz sample using the minimum of the lower ribs and abdomen as the lower signal, and similarly the maximum of the SCM and upper ribs as the upper signal:

$$opRAM(t) = \frac{max({}^{e}a_{SCM}(t), {}^{e}a_{UR}(t))}{min({}^{e}a_{LR}(t), {}^{e}a_{Abd}(t))} \tag{6.5}$$

Given that these metrics are understudied in a quantitative context, the nature of the relationships are difficult to predetermine. A common transformation for extracting information from ratio quantities is to use the logarithm of the actual ratio to compare signals, as is done in with decibels. Whereas a linear relationship prescribes an absolute meaning to the ratio unit, the logarithm can better describe and compare the behavior of the signal when multiplicative factors have more meaning. Thus, the log transformation has the potential to extract more meaning from the comparison of the rib cage compartments.

Both versions of both metrics were examined for further stastical analysis. Each of the four combinations produced a time-series; to determine a single quantity for analysis of two-minute collection, the mean of each ratio signal was calculated over all used 1-second intervals from the rate filter.

6.2.4 Respiratory Alternans

While upper muscle recruitment is the most detectable sign, other magnitude relationships are well understood to be cause for concern. Primary muscle recruitment can be dominated by motion in either the abdomen or lower ribs, and certainly varies from individual to individual. It can even vary for an individual across measurement times and postures. For a given examination, however, alternation between the abdomen and chest wall indicates fatigue of muscles used for respiratory behavior. Respiratory alternans (RA) has been observed by physicians visually and using existing breathing approaches like RIP [41] [42], but effective, quantifiable, repeatable analysis is currently unavailable in clinical settings.

To measure fluctuations in compartment contribution, we calculated the RA ratio as the ratio between the upper ribs to the abdomen magnitude, similarly to RAM:

$$RA(t) = \frac{{}^{e}a_{LR}(t)}{{}^{e}a_{Abd}(t)} \tag{6.6}$$

The log of the ratio was again used as an additional processing step, particularly because the absolute values of the magnitudes have little meaning in the context of an magnitude-relative condition like RA. Summarizing RA activity was more difficult; as opposed to the consistency of RAM, RA is characterized by a transient effect. To this end, the mean, standard deviation, maximum, and minimum signals were initially considered for each two-minute collection. Similarly to the variance of breathing rate, a lack of confidence in the ability of standard deviation to capture the variation of interest led to the development of similar sample-difference metrics. Thus, the log RA ratio signal was downsampled to 1 Hz to match the respiratory rate signal, and first differences of multiple spacings (1, 2, and 3 seconds) were taken. Again, from each of these collections, the average, median, and the maximum were calculated and included with the other summary metrics for further analysis.

6.2.5 Other Metrics: Cough, Apnea, Synchronization

The above metrics are easy to calculate with some light signal processing given the filter mask from the mean signal, the mean rate data, and the envelopes. Each of these adds depth of understanding to the respiratory motions. Still, more information is available. Cough has been studied by multiple groups as a potential indicator of disease [8] [73] and as a pattern on its own [23]. While cough was not extensively considered under the current analysis, it would follow that a larger sensor network would capture more information about the cough and further that more data would lead to more reliable cough detection. Having synchronized recordings at multiple points facilitates easier analysis of non-typical breathing events that appear as correlated motion noise.

Similarly, the absence of any motion is highly unusual, particularly on the accelerometers. The trivial case would be that the sensor is disconnected, and no signal is being recorded. This is useful information for ensuring metric validity and alerting technicians that the sensors should be adjusted. In the event that the motionless affect is transient and present on multiple sensors, the likely answer would be apnea. Apnea is a cessation in breathing that offers information about the respiratory condition. Apneas are traditionally divided into three categories: (1) central apneas, where all breathing motion and air flow ceases, (2) obstructive apneas, where air flow ceases but breathing motion continues, and (3) mixed apneas, single events that begin as a central apnea and transition to an obstructive apnea before air flow restarts. While a motion-based system cannot detect airflow separate from motion, quantification of observable apneas offers additional information in the diagnosis process.

Apnea research has been largely motivated by the development of non-invasive measurement procedures for sleep apnea. Two categories of accelerometer-based devices have been used with success: chest-motion devices, and air vibration devices. Chest devices [69] [21] [65] [121] have been shown to be effective in detecting a lack of motion in the acceleration signals relative to typical breathing rhythms as seen by a single three-axis accelerometer. Additionally, vibration devices [20] are typically positioned either below the larynx or clipped onto the septum. While the vibration methods are certainly interesting, the chest-worn devices serve as a better view of prior work, as the current sampling rate of ARK would typically be too low to catch pitched breathing information.

Apnea and apnea-related patterns were clearly seen on accelerometer signals. The effect was quantified by looking for drops in the average acceleration difference for each dimension and sensor (similar to [121]). When many sensors observed low amplitudes, apneic intervals could clearly be seen on the underlying acceleration curves in between periods of respiratory motion. Similarly, many sensors registering high average changes indicated regular breathing motion. The process was not developed further, however, for two reasons: modeling issues and the lack of gold standard data. First, the distributed approach would require either low motion cutoffs or an apnea probability model to decide which sensors were apneic and whether the set of all sensors indicated apnea. While this would be acheivable with a physician review process as in [122], such a method would be time-intensive, especially given that we do not have flow data for the larger data set which captured the apneic patterns. The combination of these concerns led to a tabling of apnea detection, though preliminary results indicate that a distributed approach would also have merits over a conventional single-sensor approach.

Another metric of interest was quantification of paradoxical motion. We developed an algorithm to track the phase difference between the abomen and lower ribs using the phase functions from the analytic representations of accelerometer signals, following from [27]. The stability of the phase difference was meant to be a proxy for paradoxical motion, which is out-of-phase behavior by the rib cage and the diaphragm in the abdomen. Stability was quantified using the entropy of the resulting distribution of phase differences; a highly-stable pair of signals would generate a clear peak at the near-constant difference. The peak would indicate the angle of difference, though we found this to be an inconsistent measure over multiple subjects. Additionally, the distribution of differences was found to be exceptionally noisy, as a transient effect on one sensor could ruin the entire distribution calculation. The noise persisted even after filtering using the mean rate information. Again, while there is considerable reason to believe that such metrics would be possible and useful, they were not included in the current study.

6.3 Evaluating Metrics with Clinical Data

790 two-minute segments from 114 trials were collected. To begin, segments that used under 30 points after mean filtering were discarded, as they were considered too sparse to generate stable estimates of means and standard deviations. From the above calculations, we computed 35 metrics for each of the 582 remaining collections. Each quantified the respiratory condition slightly differently; for instance, maximum and average sample differences were used for the mean and RA sample variation metrics, which were both also quantified by the standard deviation of each time series.

The preliminary separability analysis here used hand-selected metrics out of the pool whose distributions over the three potential outcomes (home, acute care, ICU) seemed to offer discriminatory power for severity analysis. To aid in the finding of useful trends, distributions were visualized for trial-based organizations of the segment metrics. In addition to distributions of all segments, I looked at the mean, median, maximum, minimum, first, and last metrics recorded for each trial. For each category of metric (mean rate, rate variance, RAM, RA), a candidate was selected from one of the sets that seemed to separate the interquartile ranges. The final selection was as follows:

- Respiratory Rate: Average Rate, Max Segment
- Rate Variance: Median 1-second difference of Respiratory Rate, Min Segment
- RAM: Average log opRAM ratio, Last Segment

- RA: Average 3-second difference of RA ratio, Median Segment

Distributions for the selected metrics are shown in 6.7.

This generated four metric values for each of the 108 trials. Given that the indicator data were a multi-class ordinal ranking, a fitting statistical method was needed to determine significance of the data. For this preliminary analysis, the Kruskal-Wallis ANOVA on Ranks test was selected, as it allows significance testing of median separation for categorical data. Rather than performing ANOVA testing on the data themselves, the Kruskal-Wallis method first ranks the data across the entire sample before running one-way ANOVA analysis. While it does not innately use the ordinal information, follow-up testing with pairwise Mann-Whitney U tests (with sufficient multiple comparison adjustment via Bonferroni correction or similar) allows for a more fine-grained analysis for data sets with a significance according to the Kruskal-Wallis criteria. A rank-sum approach can be applied to multi-dimensional data inputs, where each dimension of data is ranked separately, each data point's ranks are summed, then the ANOVA process is applied to the rank sum data. While this process is naive in that metrics are equally weighted, it provides a straightforward analysis of data separation across multiple categories of data. For these analyses, Kruskal-Wallis testing was performed using the *sort* and *anova1* Matlab functions to accomodate the additional step of the rank sum.

Running the Kruskal-Wallis analysis produced a p-value of 0.0002, indicating strong significance that at least one category was significantly different. This result was generally positive, but further analysis was required to ensure that it had ordinal significance (i.e. difference between home and ICU). Pairwise Mann-Whitney testing (with Tukey-Kramer honest signficant difference correction via *multcompare* Matlab function) revealed that differences in all three pairs were statistically significant (see Table 6.1).

The goal of this simplified analysis was to demonstrate the potential of the metric data. This analysis confirmed that the metrics were indeed ordinally significant, as demonstrated by the visualizations of the ranking distributions (Figure 6.8), as well as the associated p-

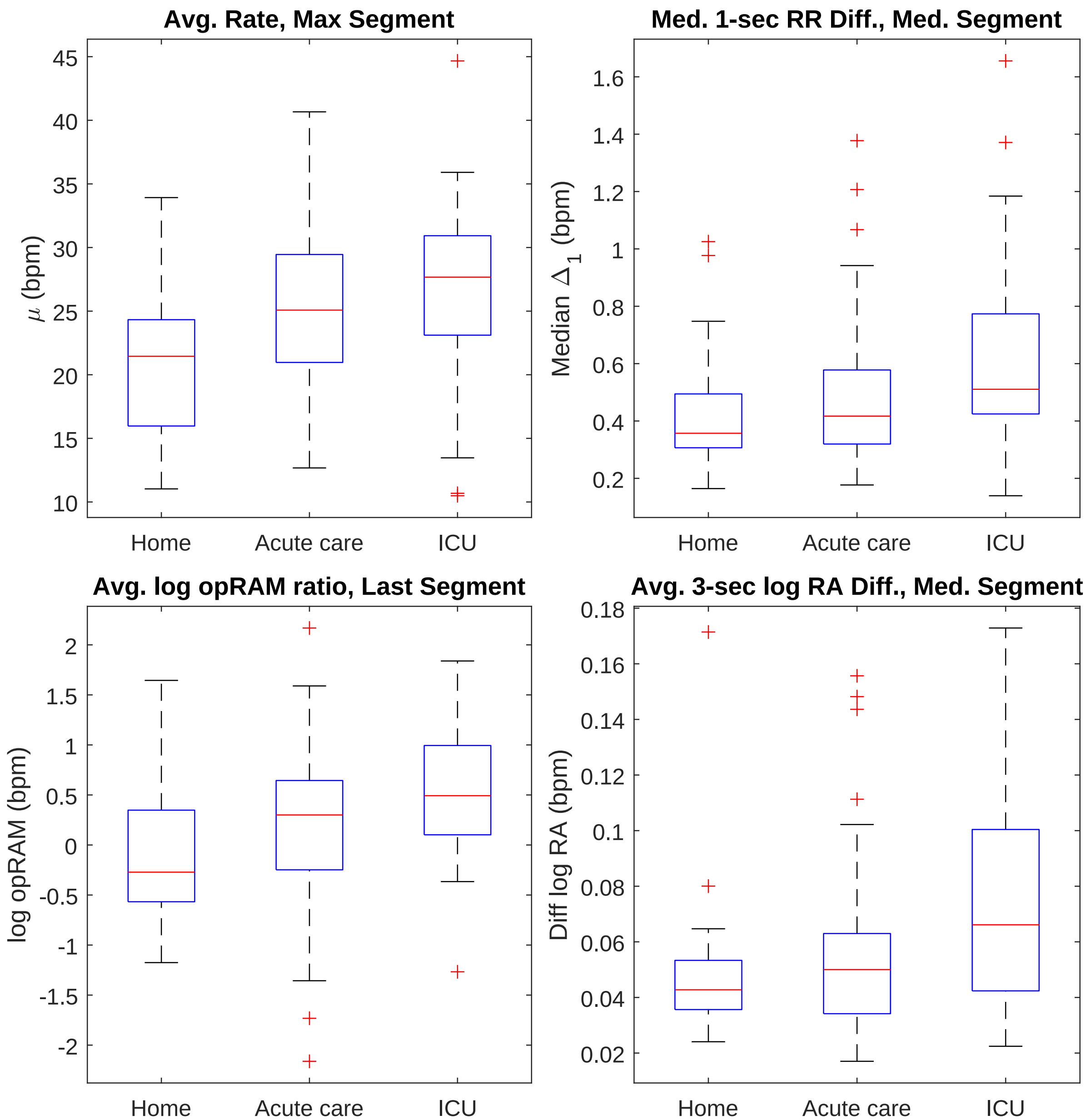


Figure 6.7: Box plots demonstrating visual separation of metric distributions by outcome. (a) Mean respiratory rate, max segment per patient. (b) Median 1-second difference of respiratory rate time series, minimum segment. (c) Average log RAM ratio, last segment. (d) Average 3-second difference of log RA ratio, median segment.

Pair	p
Home and Acute	0.0148
Home and ICU	0.0001
Acute and ICU	0.0334

Table 6.1: Mann-Whitney Pairwise Comparison Significance

values. The strong result of generalized metrics having predictive power is a crucial result for demonstrating the validity of the ARK system in a clinical context. More advanced techniques for handling metric selection and patient variability are possible and necessary for realizing the full predictive potential of ARK metrics in a clinical setting.

6.4 Ablation Study

Multiple methods have been presented for deriving the respiratory rate from a single sensor with good accuracy but with little regard for the noisy conditions in clinical and ambulatory care settings. The key advantages of a larger system like ARK are reliability through redundant measurement against an errant signal and the improved ability to detect noise that affects the whole system. The larger number of locations also allows for more complex metrics like the RAM and RA ratios, but for respiratory rate specifically it stands to reason that fewer sensors could reliably produce a respiratory rate time series. To examine the effect of reducing the number of sensors, an ablation study was performed. Specifically, the matrices of pairwise differences and cutoffs were reduced to smaller submatrices corresponding to streams from a subset of sensors. The subset started with the most useful sensor, adding sensors in order of accuracy with the caveat of only using one sensor from the upper and lower rib sensor areas until all four locations had been used. To determine the usefulness of a sensor, the number of inclusions in the rate cluster were counted for each sensor.

Results of clustering accuracy on clean data are shown in Figure 6.9a. The bilateral locations performed nearly identically, as expected. The lower ribs performed the best out of the four locations, followed by upper ribs, then abdomen, and finally SCM. The lower

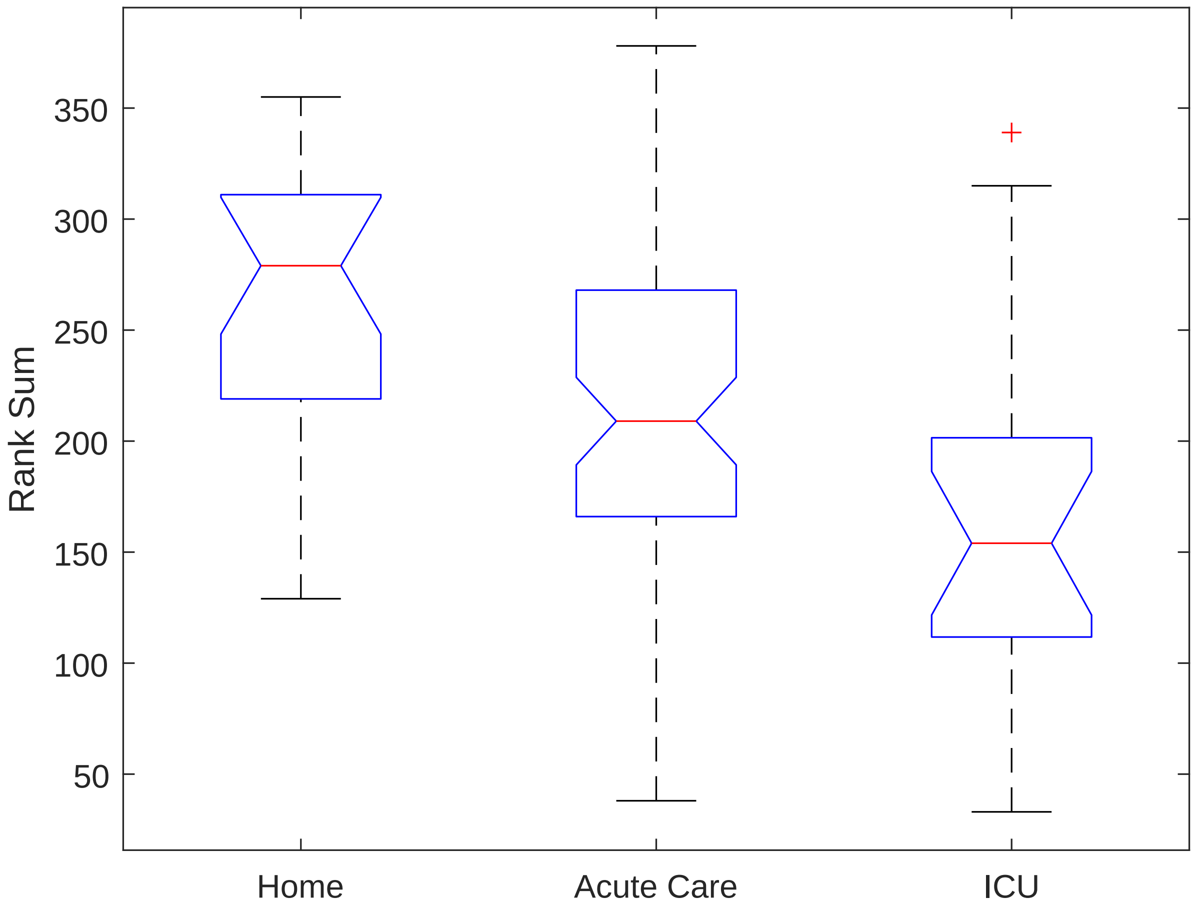
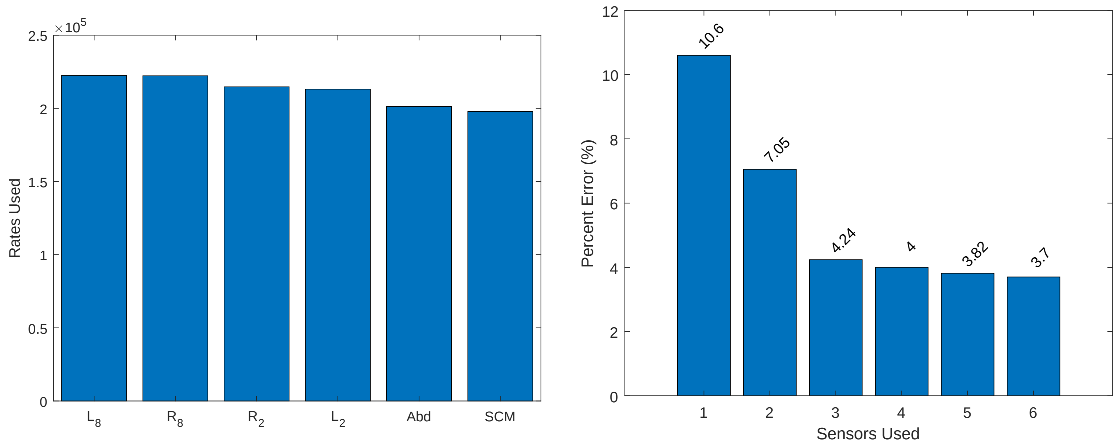


Figure 6.8: Notched box plots of the ranking distributions across different outcomes.



(a) (b)

Figure 6.9: (a) Number of inclusions in the rate cluster, by sensor (after filtering). (b) Ablation study accuracy by number of sensors used. Order of inclusion: L_8, R_2, Abdomen, SCM, R_8, L_2. Error dropped after the addition of the L_8 sensor, then again with the abdomen. Additional sensors only added marginal information, with almost no change in error.

left rib and the upper right rib sensors were slightly better than their counterpart, so they were chosen as the first and second sensors added to the test subset. The abdomen and SCM followed as third and fourth, respectively. The lower right rib outperformed the upper left rib, so it was chosen as the fifth and final subset sensor. Mean rates were recomputed for each subset, where the rates considered and resulting pairwise differences were derived exclusively from those sensors. The mean rate accuracy results on EPCL data with flow rate reference are shown in Figure 6.9b. The error rate drops after adding a second sensor and again after a third, so low that the inclusion of all three other sensors only improves by 0.5%. Clearly, rate can meaningfully be detected with only three sensors, positioned on the upper ribs, lower ribs, and abdomen. This suggests that a reduced prototype could be effective, extending ARK to other markets.

Chapter 7

Discussion

The contributions of the previous chapters provide a framework for a clinical solution to respiratory monitoring. ARK is robust to noise; algorithms are specifically designed to reject noise at all stages, while still preserving intermittent intervals of clean data. The ARK metrics are well-reasoned; each quantifies a well-known marker of respiratory distress. Respiratory rate time series have historically not been robust when produced by inertial sensors; ARK employs a well-reasoned method that can additionally produce an accurate time series variability for use in the initial studies of breath rate variability and beyond. RA has not been quantified with inertial signals; this study uses acceleration magnitudes as a proxy for the contribution of different compartments, observing variation over time. The RAM metric rounds out the group with an entirely-novel, quantitative kinematics analysis of the difficulty of breathing. These metrics have been validated in a clinical study with a prototype that is convenient and comfortable enough for extended use on patients. ARK is a contribution to the fields of respiratory kinematics measurement and clinical monitoring.

ARK is a product of the current research ecosystem. Below are commentaries on why this problem is worth pursuing with inertial sensors, key innovations driving sensor network design and application, the limitations of MEMS IMUs for kinematic sensing, and potential research directions for moving beyond this initial analysis of ARK.

7.1 Difficulties of Clinical Respiratory Sensing

The primary motivation of this work is to close the respiratory information gap in clinical settings, enabling better treatment and improving patient outcomes. Physiological measurement has been applied to every domain of medicine to facilitate data-based modeling of individual cases, but new technologies continually allow for new opportunities to improve accuracy, patient experience, and cost efficacy. ARK represents the latest development in the application of inertial sensors in clinical medicine.

A driving motivation for ARK is the knowledge of predictive signals like RAM and RA, the lack of physician time for observing these phenomena, and a lack of quantitative tools built for clinical environments. Physicians could likely predict respiratory decompensation episodes if given the time to sit bedside and carefully observe the emergence of warning signs. Unfortunately, doctors often have less than 10 minutes per patient for observation, relying on a steady stream of measurements by nurses and test results from specialists to inform their diagnostic decisions. Even if doctors did have the time to watch every patient at all times during their shift, care of patients would have to be transferred eventually; all observations would have to be transcribed onto the chart, and all qualitative assessments would be subject to interpretation. More practically, illnesses like COVID-19 drive home the importance of remote monitoring in the midst of less frequent patient observations by physicians [123] [124] [125]. All of these issues raise a single question: do monitoring methods allow for the objective description of respiratory condition with high enough quality to enable robust and thorough evaluation through routine quantitative measurements?

This work presents an affirmative argument, using inertial sensors to produce not only a respiratory rate measurement, but a rate time series enabling assessment of variation, an assessment of signal quality, and multi-modal metrics using the filtered information. We demonstrate that there is considerable discriminatory power when the metrics exclusively derived from ARK are used to characterize a patient. Further, we demonstrate that the clinical environment is not too noisy for inertial sensing to produce these results if proper care is taken to treat and clean the data; physician time is not needed to record the signals, but the results would be available when they needed them. All of this is possible with a prototype that could be easily applied in chaotic ICU and ambulatory care settings, requiring only the application of quarter-sized stickers and sensors. Inertial sensing presents one of the best cost-benefit tradeoffs for clinical sensing.

Looking back at heart rate variability underscores the importance of a lightweight, non-invasive monitoring solution for respiratory system dynamics. The development of Holter monitors for continuous, long-term, portable ECG monitoring led to a wealth of discoveries of warning signs in the form of rare rhythms and signal abnormalities [79]. Practicing physicians have noticed all kinds of subtle variations in heart rate patterns, using ECGs, processed RR interval series, and other creative solutions for monitoring. The success of monitoring devices in medicine is specifically driven by the intuition of a clinician, both what to record and how to record it. The Holter monitor allowed more physicians access to more data, and the result was a diverse panel of heart rate anomalies that herald dysfunction and deterioration. The respiratory system is a similarly chaotic process; ARK offers a potential solution for the observability problem of respiratory kinematics, giving physicians the tools to aggregate massive data sets across populations, study abnormal respiratory patterns, and design new measures for clinical respiratory practice.

7.2 Key Technologies for ARK

ARK is enabled by a set of technologies, allowing for low-cost, compact, robust sensing. The primary technological development that has driven the latest wave of respiratory kinematics methods is the proliferation of MEMS devices. New processes for commercial fabrication of sensors have allowed for multi-axis measurement of multiple inertial signals on a single, tiny chip. Further, due to semiconductor area cost gains related to Moore's Law, fully featured MEMS IMUs come with a digital processor while still maintaining a millimeter-sized die. Lastly, because the fabrication process techniques directly draw from existing circuitry methods, the chips are reasonably priced under $10. Inertial sensor networks are being proposed because MEMS has made it feasible to do so.

One novel technique in our work is the use of the Hilbert transform in concert with the Berger method to describe dynamic effects in respiratory rate-linked data like acceleration signals. As seen in the examples presented in Chapter 4, high variability in signal peaking complicates tuning of landmark-finding algorithms. The Hilbert transform is a perfect candidate for tracking these extrema and the instantaneous rate, providing multiple avenues for data analysis. Though we have used the Berger method for interpolation, the exact instantaneous frequency series could be derived from the analytic representation [75], though it would need a similar smoothing procedure to the Berger method. The information is there, and the analytic representation provides a convenient way to access it. It should be noted that the Hilbert-Berger pair isn't just useful for acceleration signals; it could be used on electromyography, skin strain, fiber optic, RIP, PPG, and air flow sensor recordings to track the notable landmarks. While these wouldn't necessarily have the ability to filter noise with redundant recordings, the Hilbert-Berger interpolation technique would still offer a generalized, robust method for deriving a highly-resolved, accurate respiratory rate time series.

Another component simplifying ARK system design was the communication interfaces provided by modern microprocessors and instrumentation platforms. Processors are often marketed with variants that include integrated communication mechanisms, like a USB or Bluetooth physical interface. While development is often slightly idiosyncratic, code examples can typically be expanded to allow for more complex functionalities with limited thought put towards the inner workings of the communication layers. This integration was crucial for the implementation of the synchronization algorithms, as any appreciable overhead to communications processing would have worsened results. Further, a robust communications interface was necessary when designing prototypes with seven nodes. USB can naturally handle many devices as a simple extension of a single device (provided bandwidth is properly distributed), but BLE generally has a limit dependent on the processing power of the specific controlling radio. We were able to run all seven devices through a single host application on an Android phone; while we encountered issues with simultaneous transmission, we found that issuing all commands sequentially was effective and feasible due to the low overhead as well as the quick turnaround between sensors. These details would have been more problematic ten years ago, but were found to be tractable in the present analysis.

Lastly, we have hand-selected features and employed statistical testing to demonstrate the power of the ARK system and metrics, avoiding the issue of interpretability that has consumed health applications of artificial intelligence (AI). Interpretability has come a long way due to active research, but there is still a lack of consensus about robust solutions. Implementation of an AI algorithm raises questions as to where the results are coming from, and these issues will hold weight until generalized models for interpretability are constructed and implemented. Added to the uncertainty of measuring novel quantities like acceleration RAM and RA ratios, uninterpretable AI presents more issues than it solves. Still, the field has come a long way since the genesis of this project, and new interpretable methods could emerge that provide understandability with all the accuracy of a deep classifier. The idea should certainly be revisited.

7.3 On Calibration

Ideally, the final results and analyses would have also examined the position of the chest, allowing for more direct application of techniques from OEP and RIP studies. Multiple sensor fusion solutions [90] [88] [89] were implemented to convert the 9-dimensional sensor-frame signals into earth-frame accelerations and positions, but the models to account for all sensor and environmental errors while still retaining accuracy for centimeter-scale deviations in position proved to be too complicated to be tractable. Specifically, MEMS accelerometers and gyroscopes are specified to have biases as large as the acceleration signals here. While the errors of accelerometers can be calibrated against to some extent, gyroscopes are notorious for the drift of their bias, necessitating online calibration methods for any sensor fusion. Magnetometers encounter both sensor and local bias interference; biases for some magnetometers are specified to be higher than the earth's magnetic field. Further, local errors can arise from electromagnetic equipment and iron distortions that prevent the tracking of a known reference vector, which is necessary for a full alignment of all sensors in the same inertial frame.

Together, these errors trip up any online process attempting to solve for noise and bias across all three sensors. The case could be made that a calibration procedure combined with accelerometer and magnetometer bias stability assumptions could reduce the online requirements to tracking gyroscope bias and drift. That being said, the accuracy of a calibration process performed over seven sensors by a technician with little embedded systems understanding would be non-robust. Any calibration setup would need to be portable and ideally work without a guaranteed flat surface. Moreover, the idea of a nurse with a dozen patients to help calibrating seven sensors for each patient needing a measurement seems like a tough sell for nurses and administrators alike. If any calibration is to happen for a system as large as ARK, it must be fully online.

Position tracking is not out of the question for future works. It would require better sensors and specifically designed sensor fusion algorithms that take advantage of the recording of the same physical system and environment. If the biases on all sensors were reduced, fusion filters like a Kalman filter would have to do less work to separate the bias errors from noise and signal. Preprocessing could be done to derive a single local magnetometer bias from multiple sensors before trying to track the vector in each frame. The final sensor fusion step would only have to account for a reduced gyroscope drift beyond the typical noise, which would also likely be lessened by improved fabrication processes. Importantly, the present work demonstrates that while position is the intuitive signal of interest, sensor-frame results can be robust, accurate, and sufficient for outcome prediction without any calibration procedure.

7.4 Recommendations for Future Works

This work demonstrates the applicability of wearable inertial sensors for problems of clinical respiratory modeling. Still, there are many opportunities to extend the results to derive more respiratory information, improve accuracy and robustness, and expand the use cases of ARK for at-home and athletic monitoring.

First, the technology should continue to be explored. It would be useful to benchmark the above methods over different postures, breathing maneuvers (e.g. forced expiration tests), and larger data sets covering more body types, respiratory illnesses, and longer recordings with continuous monitoring and regular metric updates. These are necessary to understand the most effective ways to deploy this technology in a clinical setting, but each comes with technical costs that must be negotiated. Similarly, the position methods described in the previous section should be revisited as wearable technology continues to see improvements in sensor design.

The metrics derived in this analysis relied on a single filtering process, which could easily be separated and iterated to produce better RAM and RA metrics, as well as synchrony, cough, and apnea analysis. Currently, the agreement across all rates determines good and bad data, but the information about which signals consistently produce useful rates would help design sensor-specific filters. For instance, if the lower rib and abdomen sensors are used throughout a collection interval with little help from the upper sensors, then the RA ratio could be given a high confidence while RAM would be treated with less confidence. Further, our analysis demonstrates the importance of understanding noise; apneic intervals were often categorized as noise or registered as a low-rate section of breathing, but more intelligent models could use the frequency and duration of apnea as a marker for study. Similarly, cough and paradoxical motion are both useful features that should be investigated and identified with robust methods.

Our analysis produced not only summary metrics, but underlying time series as well. The highly-comparative time series method [91] is just one example of the time series data mining that is possible, but it suggests that there exist many more metrics than those presented here that could be useful. One particular aspect is the predictability of breathing patterns and the potential for chaos theory to describe the complex physiology underlying the variation in breathing rate and depth. Future work should build on existing studies to determine which dynamic effects appear on both short and long recordings that may elude more averaged analysis.

Wearable sensors could also use some work beyond improvements in MEMS. Dedicated Medical Body Area Network (MBAN) bandwidth in the 2.3 GHz band has been reserved for medical device communication, and radios are being developed specifically for wearable applications. BLE occupies a wide band in the 2.4 GHz range, competing with WIFI and other signals with reductions in signal quality in busy environments. Other alternatives will surely become possible and popular, hopefully offering robust methods for dense sensor network communication. Further, sensors do not need to be limited to a signal sensing modality. The addition of an electrode would allow simultaneous recording of ECG and EMG signals, which would allow for cardiopulmonary coupling analysis and recording of respiratory kinematics muscular response. Cardiopulmonary integration like [69] could also validate the use of wearable seismocardiography [83] in concert with the ARK system. Seismocardiography can be used to derive a respiratory effort signal, which would allow for a duplicate respiratory signal to be derived and analyzed.

Lastly, systems for classifying metrics and determining risk scores or probabilities are necessary to begin the discussion of how best to use these metrics to improve patient outcomes. We are currently working with clustering methods to understand the plethora of metrics we have derived and the signatures of physiological behavior that we have recorded. This is one potential option for understanding and predicting signatures of distress, but more data would allow for deeper studies into the interrelation of metrics and specific respiratory events, like breathing assistance (e.g. oxygen or continuous positive airway pressure (CPAP) mask). With more data, and likely more metrics, methods would have to be adept at finding patterns in the high-dimensional data. One avenue for bridging the gap between physician knowledge and technological development is the recent work in risk score optimality [126], using artificial intelligence to learn the optimal weighting scheme for metric inputs with a model that is easy to explain and use. This modeling technique has been applied with success to patient modeling problems such as ICU risk assessment from electroencephalography (EEG) signals [127], but it would require a well-validated technique like ARK for deriving the necessary quantitative data in a clinical setting. The functional time series approach offers another method, where the instantaneous time series of metrics are used as indicators, rather than summary statistics. Concert methods consisting of multiple layers of analysis (i.e. metric extraction, time series analysis, multi-faceted classifier) would be possible for bringing the above approaches together.

This work represents one of the first steps towards validating new inertial hardware for use in effective, highly data-dependent diagnostic models; our efforts have focused specifically on robustness in clinical application. Further effort is needed to fully realize the potential of clinical respiratory sensing using inertial platforms.

Chapter 8

Conclusion

This work makes the case that inertial sensing can be robustly applied to clinical settings to derive respiratory information that has predictive power for patients with respiratory distress. Specifically, the work contributes a system for synchronized distributed breathing sensing using a novel arrangement of sensors, a robust process to derive instantaneous respiratory rates from any oscillating signal, a method for filtering the synchronized kinematics information from a network of inertial sensors, and novel methods for extracting filtered metric information about the relative magnitudes of motion from different locations on the chest wall.

These metrics quantify notions that are taught to physicians, allowing for the reliability of comparisons between metrics taken for different patients or metrics taken at different times on the same patient. Additionally, ARK shifts the burden of observation from a physician who must be at the bedside to a nurse or technician who can take the recordings with their more regular interactions. This allows a physician to review the records at will and gain temporal understanding when making diagnostic decisions, regardless of when the clinician first observed the patient. Further, ARK has been successfully used by research managers and medical students; it does not require any elaborate calibration and the training process is quick and straightforward. The lightweight design was a crucial step towards the system-level robustness.

ARK is a natural step in the progression of respiratory monitoring technology, keying into multiple trends for smart health applications. The big data phenomenon is a driving force in today's health technology ecosystem [128], taking advantage of advances in processing power on all scales. Smaller, low-power devices with arrays of sensors are now capable of generating handfuls of accurate, useful data, while newer algorithms have kept up with increases in processing power. The proliferation of MEMS devices into different areas of medicine (and beyond) will continue for the foreseeable future, driving new advancements as the technology drops below $1 per 9-axis sensor. Applications like the auxiliary use of an accelerometer for imaging gating to avoid motion noise will become more common, and systems like ARK that are dependent on small, cheap sensors could grow in size. It's not unthinkable to have more than six points on the chest wall, allowing for more complex modeling of different body types and ultimately improving accuracy and robustness. Just as systems will grow to use cheaper technology, the applications of ARK outside of the clinical setting will grow. Athletics and exercise physiology are obvious use cases, but monitoring designs for asthma patients, nursing homes, meditation, and even veterinary work are also possible.

For the above reasons, ARK, ARK filtering, and the associated metrics represent a novel and important contribution to the field of clinical respiratory sensing, and a step towards ubiquitous computing for respiratory health applications.

Bibliography

[1] MATLAB, *version 9.10.0 (R2021a)*. Natick, Massachusetts: The MathWorks Inc., 2021.

[2] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2021.

[3] RStudio Team, *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, PBC., 2020.

[4] D. Datta, *blandr: a Bland-Altman Method Comparison package for R*, 2017.

[5] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.

[6] American Thoracic Society/European Respiratory Society, "ATS/ERS Statement on respiratory muscle testing," *American Journal of Respiratory and Critical Care Medicine*, vol. 166, no. 4, pp. 518–624, Aug. 2002.

[7] 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," *Journal of Clinical Sleep Medicine : JCSM : Official Publication of the American Academy of Sleep Medicine*, vol. 8, no. 5, pp. 597–619, Oct. 2012.

[8] X. Ni, W. Ouyang, H. Jeong, J.-T. Kim, A. Tzaveils, A. Mirzazadeh, C. Wu, J. Y. Lee, M. Keller, C. K. Mummidisetty, M. Patel, N. Shawen, J. Huang, H. Chen, S. Ravi, J.-K. Chang, K. Lee, Y. Wu, F. Lie, Y. J. Kang, J. U. Kim, L. P. Chamorro, A. R. Banks, A. Bharat, A. Jayaraman, S. Xu, and J. A. Rogers, "Automated, multiparametric monitoring of respiratory biomarkers and vital signs in clinical and home settings for COVID-19 patients," *Proceedings of the National Academy of Sciences*, vol. 118, no. 19, p. e2026610118, May 2021, publisher: Proceedings of the National Academy of Sciences.

[9] S. Magder, "Bench-to-bedside review: ventilatory abnormalities in sepsis," *Critical Care (London, England)*, vol. 13, no. 1, p. 202, 2009.

[10] C. Massaroni, A. Nicolò, D. Lo Presti, M. Sacchetti, S. Silvestri, and E. Schena, "Contact-based methods for measuring respiratory rate," *Sensors (Basel, Switzerland)*, vol. 19, no. 4, p. 908, Feb. 2019.

[11] K. E. J. Philip, E. Pack, V. Cambiano, H. Rollmann, S. Weil, and J. O'Beirne, "The accuracy of respiratory rate assessment by doctors in a London teaching hospital: a cross-sectional study," *Journal of Clinical Monitoring and Computing*, vol. 29, no. 4, pp. 455–460, Aug. 2015.

[12] G. H. P. Latten, M. Spek, J. W. M. Muris, J. W. L. Cals, and P. M. Stassen, "Accuracy and interobserver-agreement of respiratory rate measurements by healthcare professionals, and its effect on the outcomes of clinical prediction/diagnostic rules," *PloS One*, vol. 14, no. 10, p. e0223155, 2019.

[13] M. J. Tobin, T. S. Chadha, G. Jenouri, S. J. Birch, H. B. Gazeroglu, and M. A. Sackner, "Breathing patterns. 1. Normal subjects," *Chest*, vol. 84, no. 2, pp. 202–205, Aug. 1983.

[14] ——, "Breathing patterns. 2. Diseased subjects," *Chest*, vol. 84, no. 3, pp. 286–294, Sep. 1983.

[15] S. R. Braun, "Respiratory Rate and Pattern," in *Clinical Methods: The History, Physical, and Laboratory Examinations*, 3rd ed., H. K. Walker, W. D. Hall, and J. W. Hurst, Eds. Boston: Butterworths, 1990.

[16] A. Verghese, B. Charlton, J. P. Kassirer, M. Ramsey, and J. P. A. Ioannidis, "Inadequacies of physical examination as a cause of medical errors and adverse events: a collection of vignettes," *The American Journal of Medicine*, vol. 128, no. 12, pp. 1322–1324.e3, Dec. 2015.

[17] A. Tulaimat, R. M. Gueret, M. F. Wisniewski, and J. Samuel, "Association between rating of respiratory distress and vital signs, severity of illness, intubation, and mortality in acutely ill subjects," *Respiratory Care*, vol. 59, no. 9, pp. 1338–1344, Sep. 2014.

[18] A. Tulaimat and W. E. Trick, "DiapHRaGM: A mnemonic to describe the work of breathing in patients with respiratory failure," *PloS One*, vol. 12, no. 7, p. e0179641, 2017.

[19] F. Schipper, R. J. G. van Sloun, A. Grassi, R. Derkx, S. Overeem, and P. Fonseca, "Estimation of respiratory rate and effort from a chest-worn accelerometer using constrained and recursive principal component analysis," *Physiological Measurement*, vol. 42, no. 4, May 2021.

[20] D. S. Morillo, J. L. R. Ojeda, L. F. C. Foix, and A. L. Jiménez, "An accelerometer-based device for sleep apnea screening," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 491–499, Mar. 2010, conference Name: IEEE Transactions on Information Technology in Biomedicine.

[21] A. H. Yüzer, H. Sümbül, and K. Polat, "A Novel Wearable Real-Time Sleep Apnea Detection System Based on the Acceleration Sensor," *IRBM*, vol. 41, no. 1, pp. 39–47, Feb. 2020.

[22] A. Cesareo, S. Gandolfi, I. Pini, E. Biffi, G. Reni, and A. Aliverti, "A novel, low cost, wearable contact-based device for breathing frequency monitoring," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, vol. 2017, pp. 2402–2405, Jul. 2017.

[23] A. Gaidhani, K. S. Moon, Y. Ozturk, S. Q. Lee, and W. Youm, "Extraction and Analysis of Respiratory Motion Using Wearable Inertial Sensor System during Trunk Motion," *Sensors (Basel, Switzerland)*, vol. 17, no. 12, Dec. 2017.

[24] C. Mottram, *Ruppel's Manual of Pulmonary Function Testing*, 11th ed. St. Louis, Missouri: Mosby, Mar. 2017.

[25] H. Jeong, J. A. Rogers, and S. Xu, "Continuous on-body sensing for the COVID-19 pandemic: Gaps and opportunities," *Science Advances*, vol. 6, no. 36, Sep. 2020.

[26] M. Chu, T. Nguyen, V. Pandey, Y. Zhou, H. N. Pham, R. Bar-Yoseph, S. Radom-Aizik, R. Jain, D. M. Cooper, and M. Khine, "Respiration rate and volume measurements using wearable strain sensors," *npj Digital Medicine*, vol. 2, no. 1, p. 8, Feb. 2019, number: 1 Publisher: Nature Publishing Group.

[27] W. B. Ashe, S. E. Innis, J. N. Shanno, C. J. Hochheimer, R. D. Williams, S. J. Ratcliffe, J. R. Moorman, and S. M. Gadrey, "Analysis of respiratory kinematics: a method to characterize breaths from motion signals," *Physiological Measurement*, vol. 43, no. 1, Jan. 2022, publisher: IOP Publishing.

[28] P. Laveneziana, A. Albuquerque, A. Aliverti, T. Babb, E. Barreiro, M. Dres, B.-P. Dubé, B. Fauroux, J. Gea, J. A. Guenette, A. L. Hudson, H.-J. Kabitz, F. Laghi, D. Langer, Y.-M. Luo, J. A. Neder, D. O'Donnell, M. I. Polkey, R. A. Rabinovich, A. Rossi, F. Series, T. Similowski, C. M. Spengler, I. Vogiatzis, and S. Verges, "ERS statement on respiratory muscle testing at rest and during exercise," *The European Respiratory Journal*, vol. 53, no. 6, p. 1801214, Jun. 2019.

[29] P. Sukul, J. K. Schubert, K. Zanaty, P. Trefz, A. Sinha, S. Kamysek, and W. Miekisch, "Exhaled breath compositions under varying respiratory rhythms reflects ventilatory variations: translating breathomics towards respiratory medicine," *Scientific Reports*, vol. 10, no. 1, p. 14109, Aug. 2020.

[30] P. Sukul, J. Bartels, P. Fuchs, P. Trefz, R. Remy, L. Rührmund, S. Kamysek, J. K. Schubert, and W. Miekisch, "Effects of COVID-19 protective face-masks and wearing durations onto respiratory-haemodynamic physiology and exhaled breath constituents," *The European Respiratory Journal*, p. 2200009, Feb. 2022.

[31] R. Schubert, H. Schwoebel, A. Mau-Moeller, M. Behrens, P. Fuchs, M. Sklorz, J. K. Schubert, S. Bruhn, and W. Miekisch, "Metabolic monitoring and assessment of anaerobic threshold by means of breath biomarkers," *Metabolomics*, vol. 8, no. 6, pp. 1069–1080, Dec. 2012.

[32] K. Konno and J. Mead, "Measurement of the separate volume changes of rib cage and abdomen during breathing," *Journal of Applied Physiology*, vol. 22, no. 3, pp. 407–422, Mar. 1967.

[33] T. J. Hixon, M. D. Goldman, and J. Mead, "Kinematics of the Chest Wall during Speech Production: Volume Displacements of the Rib Cage, Abdomen, and Lung," *Journal of Speech and Hearing Research*, vol. 16, no. 1, pp. 78–115, Mar. 1973, publisher: American Speech-Language-Hearing Association.

[34] P. J. Watson and T. J. Hixon, "Respiratory kinematics in classical (opera) singers," *Journal of Speech and Hearing Research*, vol. 28, no. 1, pp. 104–122, Mar. 1985.

[35] P. J. Watson, T. J. Hixon, E. T. Stathopoulos, and D. R. Sullivan, "Respiratory kinematics in female classical singers," *Journal of Voice*, vol. 4, no. 2, pp. 120–128, Jan. 1990.

[36] J. DalĽAva-Santucci and A. Armanganidis, "Respiratory Inductive Plethysmography," in *Pulmonary Function in Mechanically Ventilated Patients*, ser. Update in Intensive Care and Emergency Medicine, S. Benito and A. Net, Eds. Berlin, Heidelberg: Springer, 1991, pp. 121–142.

[37] M. A. Cohn, H. Watson, R. Weisshaut, F. D. Stott, and M. A. Sackner, "A transducer of noninvasive monitoring of respiration," in *Proceedings of the second international symposium on ambulatory monitoring*. London: Academic Press, 1978, pp. 119–128.

[38] M. A. Cohn, A. S. Rao, M. Broudy, S. Birch, H. Watson, N. Atkins, B. Davis, F. D. Stott, and M. A. Sackner, "The respiratory inductive plethysmograph: a new non-invasive monitor of respiration," *Bulletin Europeen De Physiopathologie Respiratoire*, vol. 18, no. 4, pp. 643–658, Aug. 1982.

[39] J. D. Sackner, A. J. Nixon, B. Davis, N. Atkins, and M. A. Sackner, "Non-invasive measurement of ventilation during exercise using a respiratory inductive plethysmograph. I," *The American Review of Respiratory Disease*, vol. 122, no. 6, pp. 867–871, Dec. 1980.

[40] C. F. Clarenbach, O. Senn, T. Brack, M. Kohler, and K. E. Bloch, "Monitoring of Ventilation During Exercise by a Portable Respiratory Inductive Plethysmograph," *Chest*, vol. 128, no. 3, pp. 1282–1290, Sep. 2005.

[41] B. P. Krieger and P. Ershowsky, "Noninvasive Detection of Respiratory Failure in the Intensive Care Unit," *Chest*, vol. 94, no. 2, pp. 254–261, Aug. 1988.

[42] J. P. Needleman, L. J. Benjamin, J. A. Sykes, and T. K. Aldrich, "Breathing Patterns During Vaso-occlusive Crisis of Sickle Cell Disease," *Chest*, vol. 122, no. 1, pp. 43–46, Jul. 2002.

[43] M. J. Tobin, G. Jenouri, B. Lind, H. Watson, A. Schneider, and M. A. Sackner, "Validation of Respiratory Inductive Plethysmography in Patients with Pulmonary Disease," *CHEST*, vol. 83, no. 4, pp. 615–620, Apr. 1983.

[44] J. C. Smith and J. Mead, "Three degree of freedom description of movement of the human chest wall," *Journal of Applied Physiology (Bethesda, Md.: 1985)*, vol. 60, no. 3, pp. 928–934, Mar. 1986.

[45] H. W. Korin, R. L. Ehman, S. J. Riederer, J. P. Felmlee, and R. C. Grimm, "Respiratory kinematics of the upper abdominal organs: a quantitative study," *Magnetic Resonance in Medicine*, vol. 23, no. 1, pp. 172–178, Jan. 1992.

[46] T. Kondo, I. Kobayashi, Y. Taguchi, Y. Ohta, and N. Yanagimachi, "A dynamic analysis of chest wall motions with MRI in healthy young subjects," *Respirology*, vol. 5, no. 1, pp. 19–25, 2000.

[47] T. Kondo, I. Kobayashi, Y. Taguchi, N. Hayama, S. Tajiri, and N. Yanagimachi, "An analysis of the chest wall motions using the dynamic MRI in healthy elder subjects," *The Tokai Journal of Experimental and Clinical Medicine*, vol. 30, no. 1, pp. 15–20, Apr. 2005.

[48] S. Kiryu, S. H. Loring, Y. Mori, N. M. Rofsky, H. Hatabu, and M. Takahashi, "Quantitative analysis of the velocity and synchronicity of diaphragmatic motion: dynamic MRI in different postures," *Magnetic Resonance Imaging*, vol. 24, no. 10, pp. 1325–1332, Dec. 2006.

[49] G. Ferrigno, P. Carnevali, A. Aliverti, F. Molteni, G. Beulcke, and A. Pedotti, "Three-dimensional optical analysis of chest wall motion," *Journal of Applied Physiology*, vol. 77, no. 3, pp. 1224–1231, Sep. 1994.

[50] S. J. Cala, C. M. Kenyon, G. Ferrigno, P. Carnevali, A. Aliverti, A. Pedotti, P. T. Macklem, and D. F. Rochester, "Chest wall and lung volume estimation by optical reflectance motion analysis," *Journal of Applied Physiology (Bethesda, Md.: 1985)*, vol. 81, no. 6, pp. 2680–2689, Dec. 1996.

[51] I. Romagnoli, F. Gigliotti, A. Galarducci, B. Lanini, R. Bianchi, D. Cammelli, and G. Scano, "Chest wall kinematics and respiratory muscle action in ankylosing spondylitis patients," *European Respiratory Journal*, vol. 24, no. 3, pp. 453–460, Sep. 2004.

[52] B. Binazzi, B. Lanini, F. Gigliotti, and G. Scano, "Breathing pattern and chest wall kinematics during phonation in chronic obstructive pulmonary disease patients," *Respiration; International Review of Thoracic Diseases*, vol. 86, no. 6, pp. 462–471, 2013.

[53] C. Massaroni, “The Use of Kinematics for Pulmonary Volume Assessment,” in *Handbook of Human Motion*. Cham: Springer International Publishing, 2018, pp. 847–862.

[54] M. Fajkus, J. Nedoma, R. Martinek, V. Vasinek, H. Nazeran, and P. Siska, “A Non-Invasive Multichannel Hybrid Fiber-Optic Sensor System for Vital Sign Monitoring,” *Sensors (Basel, Switzerland)*, vol. 17, no. 1, p. 111, Jan. 2017.

[55] B. A. Reyes, N. Reljin, Y. Kong, Y. Nam, and K. H. Chon, “Tidal Volume and Instantaneous Respiration Rate Estimation using a Volumetric Surrogate Signal Acquired via a Smartphone Camera,” *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 764–777, May 2017.

[56] M. M. Rahman, M. Y. Ahmed, T. Ahmed, B. Islam, V. Nathan, K. Vatanparvar, E. Nemati, D. McCaffrey, J. Kuang, and J. A. Gao, “BreathEasy: Assessing Respiratory Diseases Using Mobile Multimodal Sensors,” in *Proceedings of the 2020 International Conference on Multimodal Interaction*. New York, NY, USA: Association for Computing Machinery, Oct. 2020, pp. 41–49.

[57] M. M. Rahman, E. Nemati, V. Nathan, and J. Kuang, “InstantRR: Instantaneous Respiratory Rate Estimation on Context-Aware Mobile Devices,” in *13th EAI International Conference on Body Area Networks*. Cham: Springer International Publishing, 2018, pp. 267–281.

[58] T. Ahmed, M. Rahman, M. Y. Ahmed, E. Nemati, M. Dinh, N. Folkman, J. Kuang, and A. Gao, “RRMonitor: A Resource-Aware End-to-End System for Continuous Monitoring of Respiration Rate Using Earbuds,” in *43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Mexico, 2021.

[59] J. Lobotková Kimijanová, K. Neumannova, Z. Halická Hirjaková, K. Buckova Suttova, D. Bzdúšková Abrahámová, F. Hlavacka, E. Zemková, and E. G, “Using of accelerometers for recording of respiratory movements during quiet stance,” in *Piešťany Physiotherapy Days 2016*, Piešťany, Slovakia, Sep. 2016.

[60] M. Jafari Tadi, T. Koivisto, M. Pänkäälä, and A. Paasio, “Accelerometer-Based Method for Extracting Respiratory and Cardiac Gating Information for Dual Gating during Nuclear Medicine Imaging,” *International Journal of Biomedical Imaging*, vol. 2014, p. e690124, Jul. 2014.

[61] P. Grossman, “The LifeShirt: a multi-function ambulatory system monitoring health, disease, and medical intervention in the real world,” *Studies in Health Technology and Informatics*, vol. 108, pp. 133–141, 2004.

[62] S. Lapi, F. Lavorini, G. Borgioli, M. Calzolai, L. Masotti, M. Pistolesi, and G. A. Fontana, “Respiratory rate assessments using a dual-accelerometer device,” *Respiratory Physiology & Neurobiology*, vol. 191, pp. 60–66, Jan. 2014.

[63] S. Hughes, H. Liu, and D. Zheng, “Influences of Sensor Placement Site and Subject Posture on Measurement of Respiratory Frequency Using Triaxial Accelerometers,” *Frontiers in Physiology*, vol. 11, p. 823, Jul. 2020.

[64] S. Saha, M. Kabir, N. Montazeri, B. Gavrilovic, K. Zhu, A. Yadollahi, and H. Alshaer, “Apnea-hypopnea index (AHI) estimation using breathing Sounds, accelerometer and pulse oximeter,” *ERJ Open Research*, vol. 5, no. suppl 3, p. 63, Apr. 2019.

[65] A. Manoni, F. Loreti, V. Radicioni, D. Pellegrino, L. Della Torre, A. Gumiero, D. Halicki, P. Palange, and F. Irrera, “A New Wearable System for Home Sleep Apnea Testing, Screening, and Classification,” *Sensors (Basel, Switzerland)*, vol. 20, no. 24, p. E7014, Dec. 2020.

[66] M. J. Tadi, J. Teuho, E. Lehtonen, A. Saraste, M. Pänkäälä, T. Koivisto, and M. Teräs, “A novel dual gating approach using joint inertial sensors: implications for cardiac PET imaging,” *Physics in Medicine & Biology*, vol. 62, no. 20, pp. 8080–8101, Oct. 2017.

[67] S. Preejith, A. Jeelani, P. Maniyar, J. Joseph, and M. Sivaprakasam, “Accelerometer based system for continuous respiratory rate monitoring,” in *2017 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, May 2017, pp. 171–176.

[68] G.-Z. Liu, Y.-W. Guo, Q.-S. Zhu, B.-Y. Huang, and L. Wang, “Estimation of Respiration Rate from Three-Dimensional Acceleration Data Based on Body Sensor Network,” *Telemedicine Journal and e-Health*, vol. 17, no. 9, pp. 705–711, Nov. 2011.

[69] K. Pandia, O. T. Inan, G. T. A. Kovacs, and L. Giovangrandi, “Extracting respiratory information from seismocardiogram signals acquired on the chest using a miniature accelerometer,” *Physiological Measurement*, vol. 33, no. 10, pp. 1643–1660, Sep. 2012.

[70] A. Cesareo, Y. Previtali, E. Biffi, and A. Aliverti, “Assessment of Breathing Parameters Using an Inertial Measurement Unit (IMU)-Based System,” *Sensors (Basel, Switzerland)*, vol. 19, no. 1, Dec. 2018.

[71] A. Cesareo, E. Biffi, D. Cuesta-Frau, M. G. D’Angelo, and A. Aliverti, “A novel acquisition platform for long-term breathing frequency monitoring based on inertial measurement units,” *Medical & Biological Engineering & Computing*, vol. 58, no. 4, pp. 785–804, Apr. 2020.

[72] K. Lee, X. Ni, J. Y. Lee, H. Arafa, D. J. Pe, S. Xu, R. Avila, M. Irie, J. H. Lee, R. L. Easterlin, D. H. Kim, H. U. Chung, O. O. Olabisi, S. Getaneh, E. Chung, M. Hill, J. Bell, H. Jang, C. Liu, J. B. Park, J. Kim, S. B. Kim, S. Mehta, M. Pharr, A. Tzavelis, J. T. Reeder, I. Huang, Y. Deng, Z. Xie, C. R. Davies, Y. Huang, and J. A. Rogers, “Mechanoacoustic sensing of physiological processes and body motions via a soft wireless device placed at the suprasternal notch,” *Nature biomedical engineering*, vol. 4, no. 2, pp. 148–158, Feb. 2020.

[73] L. Lonini, N. Shawen, O. Botonis, M. Fanton, C. Jayaraman, C. K. Mummidisetty, S. Y. Shin, C. Rushin, S. Jenz, S. Xu, J. A. Rogers, and A. Jayaraman, "Rapid Screening of Physiological Changes Associated With COVID-19 Using Soft-Wearables and Structured Activities: A Pilot Study," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 9, pp. 1–11, 2021.

[74] P. Dehkordi, A. Garde, B. Molavi, J. M. Ansermino, and G. A. Dumont, "Extracting Instantaneous Respiratory Rate From Multiple Photoplethysmogram Respiratory-Induced Variations," *Frontiers in Physiology*, vol. 9, p. 948, 2018.

[75] A. Taboada-Crispi, L. Rivera, and M. B. Pérez, "Algorithms to estimate the instantaneous-frequency of a respiratory time-varying sequence," *Revista Cubana de Ciencias Informáticas*, vol. 14, no. 4, pp. 102–122, Dec. 2020.

[76] F.-T. Wang, H.-L. Chan, C.-L. Wang, H.-M. Jian, and S.-H. Lin, "Instantaneous Respiratory Estimation from Thoracic Impedance by Empirical Mode Decomposition," *Sensors (Basel, Switzerland)*, vol. 15, no. 7, pp. 16 372–16 387, Jul. 2015.

[77] L. Ren, H. Wang, K. Naishadham, Q. Liu, and A. E. Fathy, "Non-invasive detection of cardiac and respiratory rates from stepped frequency continuous wave radar measurements using the state space method," in *2015 IEEE MTT-S International Microwave Symposium*, May 2015, pp. 1–4.

[78] L. Mirmohamadsadeghi, S. Fallet, V. Moser, F. Braun, and J.-M. Vesin, "Real-time respiratory rate estimation using imaging photoplethysmography inter-beat intervals," in *2016 Computing in Cardiology Conference (CinC)*, Sep. 2016, pp. 861–864.

[79] G. Billman, "Heart Rate Variability – A Historical Perspective," *Frontiers in Physiology*, vol. 2, p. 86, 2011.

[80] L. Mirmohamadsadeghi and J.-M. Vesin, "Respiratory rate estimation from the ECG using an instantaneous frequency tracking algorithm," *Biomedical Signal Processing and Control*, vol. 14, pp. 66–72, Nov. 2014.

[81] V. Hartmann, H. Liu, F. Chen, W. Hong, S. Hughes, and D. Zheng, "Toward Accurate Extraction of Respiratory Frequency From the Photoplethysmogram: Effect of Measurement Site," *Frontiers in Physiology*, vol. 10, 2019.

[82] E. J. W. Maarsingh, L. A. van Eykern, A. B. Sprikkelman, M. O. Hoekstra, and W. M. C. van Aalderen, "Respiratory muscle activity measured with a noninvasive EMG technique: technical aspects and reproducibility," *Journal of Applied Physiology*, vol. 88, no. 6, pp. 1955–1961, Jun. 2000, publisher: American Physiological Society.

[83] P. Castiglioni, A. Faini, G. Parati, and M. Di Rienzo, "Wearable seismocardiography," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2007, 2007, pp. 3954–3957.

[84] D. Garrido, J. J. Assioun, A. Keshishyan, M. A. Sanchez-Gonzalez, and B. Goubran, "Respiratory Rate Variability as a Prognostic Factor in Hospitalized Patients Transferred to the Intensive Care Unit," *Cureus*, vol. 10, no. 1, p. e2100, 2018.

[85] R. Soni and M. Muniyandi, "Breath Rate Variability: A Novel Measure to Study the Meditation Effects," *International Journal of Yoga*, vol. 12, no. 1, pp. 45–54, 2019.

[86] C. A. Cohen, G. Zagelbaum, D. Gross, C. Roussos, and P. T. Macklem, "Clinical manifestations of inspiratory muscle fatigue," *The American Journal of Medicine*, vol. 73, no. 3, pp. 308–316, Sep. 1982.

[87] D. K. Shaeffer, "MEMS inertial sensors: A tutorial overview," *IEEE Communications Magazine*, vol. 51, no. 4, pp. 100–109, Apr. 2013.

[88] A. Sabatini, "Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 7, pp. 1346–1356, Jul. 2006.

[89] A. M. Sabatini, "Kalman-filter-based orientation determination using inertial/magnetic sensors: observability analysis and performance evaluation," *Sensors (Basel, Switzerland)*, vol. 11, no. 10, pp. 9182–9206, 2011.

[90] S. O. H. Madgwick, A. J. L. Harrison, and R. Vaidyanathan, "Estimation of IMU and MARG orientation using a gradient descent algorithm," in *2011 IEEE International Conference on Rehabilitation Robotics*, Jun. 2011, pp. 1–7.

[91] B. D. Fulcher and N. S. Jones, "Highly Comparative Feature-Based Time-Series Classification," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 12, pp. 3026–3037, Dec. 2014.

[92] M. Pagani, F. Lombardi, S. Guzzetti, O. Rimoldi, R. Furlan, P. Pizzinelli, G. Sandrone, G. Malfatto, S. Dell'Orto, and E. Piccaluga, "Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympatho-vagal interaction in man and conscious dog," *Circulation Research*, vol. 59, no. 2, pp. 178–193, Aug. 1986.

[93] T. F. of the European Society of Cardiology the North American Society of Pacing, "Heart Rate Variability," *Circulation*, vol. 93, no. 5, pp. 1043–1065, Mar. 1996.

[94] S. Craft, "Achieving Residential Demand Response through Predictive Evaluation of Thermal Response (ETHER) Models: An ARIMA Approach," Ph.D. dissertation, University of Virginia, Apr. 2012.

[95] J. A. Stuchi, L. Boccato, and R. Attux, "Frequency learning for image classification," *arXiv (2006.15476 [cs])*, Jun. 2020.

[96] B. Logan, "Mel Frequency Cepstral Coefficients for Music Modeling," *ISMIR 2000*, 2000.

[97] J. W. Denver, S. F. Reed, and S. W. Porges, "Methodological Issues in the Quantification of Respiratory Sinus Arrhythmia," *Biological psychology*, vol. 74, no. 2, pp. 286–294, Feb. 2007.

[98] F. W. King, *Hilbert Transforms*, ser. Encyclopedia of Mathematics and its Applications. Cambridge: Cambridge University Press, 2009, vol. 1.

[99] N. E. Huang and S. S. Shen, Eds., *Hilbert-Huang transform and its applications*, ser. Interdisciplinary mathematical sciences. Singapore ; Hackensack, NJ ; London: World Scientific, 2005, vol. 5.

[100] S. M. Pincus, I. M. Gladstone, and R. A. Ehrenkranz, "A regularity statistic for medical data analysis," *Journal of Clinical Monitoring*, vol. 7, no. 4, pp. 335–345, Oct. 1991.

[101] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, Jun. 2000.

[102] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman, "Sample entropy analysis of neonatal heart rate variability," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 283, no. 3, pp. R789–R797, Sep. 2002.

[103] P. A. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez, and J. G. Conde, "Research electronic data capture (REDCap)–a metadata-driven methodology and workflow process for providing translational research informatics support," *Journal of Biomedical Informatics*, vol. 42, no. 2, pp. 377–381, Apr. 2009.

[104] P. A. Harris, R. Taylor, B. L. Minor, V. Elliott, M. Fernandez, L. O'Neal, L. McLeod, G. Delacqua, F. Delacqua, J. Kirby, and S. N. Duda, "The REDCap consortium: Building an international community of software platform partners," *Journal of Biomedical Informatics*, vol. 95, p. 103208, Jul. 2019.

[105] G. Borg, "Perceived exertion as an indicator of somatic stress," *Scandinavian Journal of Rehabilitation Medicine*, vol. 2, no. 2, pp. 92–98, 1970.

[106] R. L. Seip, D. Snead, E. F. Pierce, P. Stein, and A. Weltman, "Perceptual responses and blood lactate concentration: effect of training state," *Medicine and Science in Sports and Exercise*, vol. 23, no. 1, pp. 80–87, Jan. 1991.

[107] R. E. Kleiger, J. P. Miller, J. T. Bigger, and A. J. Moss, "Decreased heart rate variability and its association with increased mortality after acute myocardial infarction," *The American Journal of Cardiology*, vol. 59, no. 4, pp. 256–262, Feb. 1987.

[108] M. A. Woo, W. G. Stevenson, D. K. Moser, and H. R. Middlekauff, "Complex heart rate variability and serum norepinephrine levels in patients with advanced heart failure," *Journal of the American College of Cardiology*, vol. 23, no. 3, pp. 565–569, Mar. 1994.

[109] H. V. Huikuri, T. Seppänen, M. J. Koistinen, K. J. Airaksinen, M. Ikäheimo, A. Castellanos, and R. J. Myerburg, “Abnormalities in Beat-to-Beat Dynamics of Heart Rate Before the Spontaneous Onset of Life-Threatening Ventricular Tachyarrhythmias in Patients With Prior Myocardial Infarction,” *Circulation*, vol. 93, no. 10, pp. 1836–1844, May 1996.

[110] M. P. Tulppo, T. H. Makikallio, T. E. Takala, T. Seppanen, and H. V. Huikuri, “Quantitative beat-to-beat analysis of heart rate dynamics during exercise,” *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 271, no. 1, pp. H244–H252, Jul. 1996.

[111] H. V. Huikuri, T. H. Mäkikallio, C.-K. Peng, A. L. Goldberger, U. Hintze, and M. Møller, “Fractal Correlation Properties of R-R Interval Dynamics and Mortality in Patients With Depressed Left Ventricular Function After an Acute Myocardial Infarction,” *Circulation*, vol. 101, no. 1, pp. 47–53, Jan. 2000.

[112] J. T. Bigger, R. C. Steinman, L. M. Rolnitzky, J. L. Fleiss, P. Albrecht, and R. J. Cohen, “Power Law Behavior of RR-Interval Variability in Healthy Middle-Aged Persons, Patients With Recent Acute Myocardial Infarction, and Patients With Heart Transplants,” *Circulation*, vol. 93, no. 12, pp. 2142–2151, Jun. 1996.

[113] E. H. Hon and S. T. Lee, “Electronic evaluation of the fetal heart rate. VIII. patterns preceding fetal death, further observations,” *American Journal of Obstetrics and Gynecology*, vol. 87, pp. 814–826, Nov. 1963.

[114] M. M. Wolf, G. A. Varigos, D. Hunt, and J. G. Sloman, “Sinus arrhythmia in acute myocardial infarction,” *The Medical Journal of Australia*, vol. 2, no. 2, pp. 52–53, Jul. 1978.

[115] G. B. Drummond, D. Fischer, M. Lees, A. Bates, J. Mann, and D. Arvind, “Classifying signals from a wearable accelerometer device to measure respiratory rate,” *ERJ Open Research*, vol. 7, no. 2, pp. 00 681–2020, Apr. 2021.

[116] J. A. Vázquez-Segura, A. R. Mejía-Rodríguez, and B. A. Reyes, “Estimación Óptica Remota de la Actividad y Frecuencia Respiratoria durante Diversas Maniobras Respiratorias,” *Memorias del Congreso Nacional de Ingeniería Biomédica*, vol. 5, no. 1, pp. 114–117, Oct. 2018.

[117] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, “Phase synchronization of chaotic oscillators,” *Physical Review Letters*, vol. 76, no. 11, pp. 1804–1807, Mar. 1996.

[118] L. Marple, “Computing the discrete-time ”analytic” signal via FFT,” *IEEE Transactions on Signal Processing*, vol. 47, no. 9, pp. 2600–2603, Sep. 1999.

[119] R. D. Berger, S. Akselrod, D. Gordon, and R. J. Cohen, “An Efficient Algorithm for Spectral Analysis of Heart Rate Variability,” *IEEE Transactions on Biomedical Engineering*, vol. BME-33, no. 9, pp. 900–904, Sep. 1986.

[120] P. R. J. Östergård, "A fast algorithm for the maximum clique problem," *Discrete Applied Mathematics*, vol. 120, no. 1, pp. 197–207, Aug. 2002.

[121] Y.-S. Hsu, T.-Y. Chen, D. Wu, C.-M. Lin, J.-N. Juang, and W.-T. Liu, "Screening of obstructive sleep apnea in patients who snore using a patch-type device with electrocardiogram and 3-axis accelerometer," *Journal of Clinical Sleep Medicine*, vol. 16, no. 7, pp. 1149–1160, Jul. 2020.

[122] H. Lee, C. G. Rusin, D. E. Lake, M. T. Clark, L. Guin, T. J. Smoot, A. O. Paget-Brown, B. D. Vergales, J. Kattwinkel, J. R. Moorman, and J. B. Delos, "A new algorithm for detecting central apnea in neonates," *Physiological measurement*, vol. 33, no. 1, pp. 1–17, Jan. 2012.

[123] J. J. Reeves, J. W. Ayers, and C. A. Longhurst, "Telehealth in the COVID-19 Era: A Balancing Act to Avoid Harm," *Journal of Medical Internet Research*, vol. 23, no. 2, p. e24785, Feb. 2021.

[124] D. M. Gelfman, "Will the Traditional Physical Examination Be Another Casualty of COVID-19?" *The American Journal of Medicine*, vol. 134, no. 3, pp. 299–300, Mar. 2021.

[125] P. Hyman, "The Disappearance of the Primary Care Physical Examination-Losing Touch," *JAMA internal medicine*, vol. 180, no. 11, pp. 1417–1418, Nov. 2020.

[126] B. Ustun and C. Rudin, "Learning Optimized Risk Scores," *Journal of Machine Learning Research*, vol. 20, no. 150, pp. 1–75, 2019.

[127] A. F. Struck, B. Ustun, A. R. Ruiz, J. W. Lee, S. M. LaRoche, L. J. Hirsch, E. J. Gilmore, J. Vlachy, H. A. Haider, C. Rudin, and M. B. Westover, "Association of an Electroencephalography-Based Risk Score With Seizure Probability in Hospitalized Patients," *JAMA neurology*, vol. 74, no. 12, pp. 1419–1424, Dec. 2017.

[128] J. Andreu-Perez, C. C. Y. Poon, R. D. Merrifield, S. T. C. Wong, and G.-Z. Yang, "Big Data for Health," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 4, pp. 1193–1208, Jul. 2015.